Classical Conditioning Alters Short Noncoding RNA Expression in Drosophila

Classical Conditioning Alters Short Noncoding RNA Expression In *Drosophila*

A dissertation presented

by

Silas Dana Maniatis

to

The Department of Molecular and Cellular Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biochemistry

Harvard University

Cambridge, Massachusetts

April 2015

Professor Sam Kunes Silas Dana Maniatis

Classical Conditioning Alters Short Noncoding RNA Expression In *Drosophila*

Abstract

MicroRNAs (miRNAs) and other classes of short non-coding RNAs regulate essential processes in the development and function of the nervous system. Regulation of miRNAs by neural activity has also been reported. Recently, instances of piwi interacting RNA (piRNA) and endogenous short interfering RNA (esiRNA) mediated modulation of neural physiology have been reported. To better understand the role of miRNAs and other classes of short non-coding RNAs in long term memory (LTM) formation, we have conducted high throughput sequencing on 15-35nt RNAs isolated from heads of *Drosophila* that have been subjected to aversive olfactory conditioning. We developed genome wide profiles of miRNA, piRNA, and esiRNA, and tested for differential expression following conditioning. We find that 5 miRNAs exhibit significant regulation in the conditioned group. We identify several esiRNA generating loci within genes required for olfactory LTM formation. Our data reveal that an intron of the multiple wing hairs (mwh) gene forms secondary structures and generates esiRNAs following conditioning from regions that correspond to lysozyme family genes located within the mwh intron. We find that piRNAs are produced in fly heads, and that a small set of piRNA generating loci mapping to LTR retrotransposons are significantly down regulated following conditioning. In addition to the well characterized classes of short non coding RNAs, we describe a set of transcripts that produce large numbers of reads with a broad size distribution from the sense strand. We find that a subset of these are regulated following treatment and contain consensus elements that may be involved in their regulation. We investigate expression of one such gene with dramatically up-regulated reads following treatment, the *Drosophila* beta-site APP-cleaving enzyme (dBACE), and find that increased reads reflect increased mRNA levels. Further, we find that the target of dBACE protein, drosophila β amyloid protein precursor-like (APPL), is subjected to increased cleavage following conditioning, and that dBACE is required for LTM formation, but not for learning or STM.

Table of contents

Acknowledgements

The writing of this dissertation represents the final act in a significant phase of my intellectual and personal life. Though its completion obviously brings me great satisfaction, I feel I will not be able fully embrace and enjoy this milestone without first acknowledging the critical support and important contributions of those around me. I take the first step toward an adequate expression of my gratitude here, but I do so knowing that these words of thanks cannot really suffice.

I wish to thank Sam Kunes for his guidance and support during my years in his lab. He provided me with the immense latitude I needed in developing the work described in this dissertation. He continued to support me when other advisors might not have, and showed enormous patience and faith in my work. His daily presence in the fly room and at the bench made for a unique lab environment, and fostered exchanges that are unlikely to have occurred in an office setting. I am unaware of many other examples in which a senior faculty member actually performed experiments themselves that are incorporated in their students' dissertations. His many hours of labor in setting up and conducting the olfactory classical conditioning experiments included in this dissertation were absolutely vital to my work.

I also need to thank the members of my advisory committee, Craig Hunter, Venkatesh Murthy, and Joshua Sanes. They too showed enormous patience and faith in my work, and without their interventions and encouragement at key points, I would not have been able to complete my graduate studies.

My friends and colleagues from the de Bivort, Francis, Kunes, Lichtman, and Maniatis labs, as well as those from the “Secret” journal club have my gratitude for the many forms of advise and help they have given me. I thank them for keeping science fun and my interests broad.

During the course of their graduate work, many doctoral students receive significant support from their families. I realize that I cannot really compare the value of my version of such a thing to that of anyone else’s, but were there an objective scale, I am confident that mine would be at the top end. First, and foremost, my wife has stood behind me at every turn, and in every manner possible. A catalog of the ways in which her support has been essential to me as a person, and to the completion of this work would be far to long to print here. Also, the aggregate effect of the various forms of her support has been far greater than their sum. Again, words will fail, but in essence, she has provided me with loving encouragement, shepherded me through difficulties, made sure I have celebrated my successes, and helped those around us who have not been down the doctoral path to understand my struggles. She has been my cut man, my advocate, and my defender, and for all of this I am grateful in a way that cannot be concisely distilled, and so I will not try to do so here. Most critically though, while being the most loving and supportive spouse one could imagine, she has done so while simultaneously pushing me to keep progressing. This has undoubtedly been a herculean task, but she has managed to attack it with positivity and love. The patience and endurance of her loving support has kept me going through times when nothing else would have. Her remarkable strength has allowed her to do all of

this while simultaneously pursuing her own career in architecture. She has achieved more than can be listed here, but her work in many ways reflects the person she is. Her decision to change the direction and pursue architecture in itself displays her bravery. Her elegance finds expression in many of her projects, her drive manifests in the volume and level of her work, and her toughness was in evidence when she learned to weld and to operate heavy machinery. Her accomplishments and motivation are my inspiration. I will spend the rest of my life making sure she understands how much I value all that she has done for me, how much I respect her, and this brief section of my dissertation cannot contain the years of gratitude I owe her.

I also have the great fortune of being born into a family for whom the philosophy of science is a guiding force. I am the son of a man once described to me by an accomplished professor of biology as "A scientist's scientist in the same way that Ted Williams was a baseball player's baseball player." At the time, this sounded great, but I had not yet begun my graduate career, and was thus not well enough read in biology to appreciate the remark. Reflecting back upon the comment now, I realize how apt it was. Much as Williams' book "The Science of Hitting" remains a foundational text for developing baseball players, "Molecular Cloning" has allowed generations of scientists working in diverse areas of biology to bring the power of molecular biology and biochemistry to their work, and I have had the pleasure of putting this text to use in my own work. My father's course on gene regulation was required for all students in my program, and it was one of the most difficult I have taken. By relying solely on current publications for course materials, he and his co-

instructor Dr. Nicole Francis not only gave us a cutting edge education in the science of gene regulation, but also demanded that we learn how to read and evaluate the literature. The knowledge and skills I learned in this class remain amongst the most important from my graduate studies. More personally, my father has always respected and supported my choices in intellectual pursuits, be it computers, engineering, psychology, or biology, and he has always helped me demand my best from myself in these pursuits. Though these interactions have not always been the easiest, I can say with certainty that I could never have accomplished what I have without them. Such is the nature of true friendship and collegiality. I appreciate all of these facets of our relationship immensely, yet the intangible aspects of our father and son relationship are of even greater importance, and cannot be neatly encapsulated here. Recently, I sat on my couch holding my week old daughter while my father sat next to me discovering her twin brother's first game. Minutes later, and still holding my children, we were discussing the science of a novel approach to treating heart disease. I will let the special nature of this moment express the many things I have to thank my father for.

My mother is also a scientist by training, having obtained a master's in geology before I was born. She put this career on hold for our family, yet fostered my intellectual curiosity and development every bit as much as my father has. Her caring turned ER visits for stitches into lessons in the biology of wounds, and her patience recognized the educational value in my near detonation of the bathroom following after school electrolysis of water. She even gave me my own plot in the garden, and let me plant something different and of my own choosing every year.

But the most salient things she has taught me concern one's attitude towards life. Rather than try in futility to keep tabs on her teenage boys, she kept our kitchen stocked and our door open. As a result, she kept my friends and I off the street when she had to return home from law school or work later than I would get out of school. She has never hesitated to disassemble a faucet in need of repair, but also won't hesitate to give customer service a piece of her mind on the phone. After obtaining her J.D. in her middle age and building a career as a lawyer, she was diagnosed with breast cancer. Having defeated that challenge, she decided to take on a PhD in Geology in her 50s, and even lived completely alone in the wilderness of Nova Scotia for weeks on end conducting fieldwork. All the while, she has made it a point to put her family first, and to instill in her sons the vital importance of right now. The years of hard work it took to teach me these things by example have left a debt I will never be able to repay.

In addition to my biological brother, I have acquired several brothers through friendship. Together, my brothers have taught me the things in life that one's parents and education can't. They have challenged me, disciplined me, held me up in triumph, and commiserated with me in defeat. They have expanded my intellectual horizons beyond academics, and honed my philosophy of life... But these things have always been a collaboration, and I know that we will continue to form like Voltron.

Without the help of those mentioned here, and others, I would not be who I am, nor would I have completed the work contained in this dissertation. My gratitude cannot be bound by words and is eternal... Thank you.

Chapter I

Introduction

Of the advantages gained by organisms possessing a nervous system, perhaps none is more remarkable than the ability to anticipate future events based upon past experience. Learning and lasting retention of acquired knowledge provide a means by which organisms can adapt to new environmental circumstances within the lifespan of an individual. In the human context, it is arguable that learning and memory are the core components that define who we are as individuals, as societies, and as a species. As such, understanding the basic mechanisms by which animals learn and remember have long been, and remain major areas of study. The underlying physiology of learning and memory is largely composed of the formation, destruction, and modulation of the efficacy of synaptic connections, a set of processes termed synaptic plasticity. Much as memories can be trivial and fleeting, or profound and long enduring, so can be the changes in synapses, and neural circuitry associated with learning and memory. Further, with repeated presentation of information, memories that might otherwise be short lived can be made to last. The underlying molecular mechanisms are known to reflect these experiential aspects of learning and memory as well, and are largely conserved from invertebrates to humans. For example, triggering of the cyclic adenosine monophosphate (cAMP) signaling pathway by synaptic activity relevant to learning results in the modification of existing synaptic proteins, and leads to the initial and least enduring changes in synaptic efficacy. Formation of lasting memories involves transcription and protein synthesis, and can lead to enduring changes in gene regulation, synaptic efficacy, and synaptic number. Modulation of the activity of genes under the control of the cAMP response promoter element

(CRE) via the transcription factor cAMP response promoter element binding protein (CREB) appears to be a central and conserved aspect of the conversion of an initial learning experience to a stabile memory. The fact that these major mechanisms governing synaptic plasticity are evolutionarily conserved allows the study of gene regulation following long-term memory formation in invertebrates to yield insight into the most profound cognitive functions of higher organisms, including humans. Therefore, in the first part of this chapter, I will discuss the formation of aversive olfactory memory, its underpinnings in synaptic plasticity, and the role of *Drosophila* genetics in understanding the processes involved. Recently, studies of gene regulation in neural tissues have shown that a variety of short, noncoding RNAs modulate the production of factors required for neural development, synaptic plasticity, and learning and memory. Thus, in the second section of this chapter, I will discuss evidence that short noncoding RNAs regulate synaptic plasticity, and distinctions between the several classes of these RNAs. One of the most widely known neural disorders affecting memory is Alzheimer's disease (AD). A central feature of AD is the formation of synaptic plaques composed of β-amyloid precursor protein (APP) cleavage products. *Drosophila* possess a homologue of APP, the APP like protein (APPL), and recently *Drosophila* homologues of the proteases that cleave APP have been identified. Therefore, it is now possible to use the power of *Drosophila* behavioral genetics to study this disease. It has already been shown that disruptions of the regulated expression and cleavage of APPL result in synaptic abnormalities and long-term memory (LTM) defects in flies. Accordingly, in the

final section, I will provide an overview of APPL cleavage and its role in synaptic plasticity and memory.

Part I: Synaptic Plasticity In Learning And Memory, And Olfactory Memory In Drosophila

Synaptic plasticity is the physiological basis of learning and memory

For many organisms possessing a nervous system, the ability to form memories is crucial for survival. Learning the stimuli associated with a good place to find food can not only help an animal return to that location, but also to find new, similar locations where food is likely to be present. For animals spanning the evolutionary distance from mollusks to mammals, the distinctive odor of a predator can provide a crucial warning if the association of the odor with the predator can be learned and later remembered. Some birds, and many mammals are able to learn useful behaviors by observing others, and to retain the new behavior long enough to pass the behavior on to their own offspring. In this way, learning and memory are the foundation of culture. Memory in many ways defines who we are and how we react to life events, and as such, it was the focus of early work in psychology. To Freud, memories, conscious or repressed, were the root cause of psychopathologies. He and others exerted great effort toward an understanding of the formation, retention, and extinction of memories in the context of treating mental illness. Though Freud preferred empirical evidence, little was available in his time, and he

viewed observations by a neutral psychoanalyst as a valid method for understanding the inner workings of the mind. Accordingly, this early work dealt largely with the cognitive aspects of memory, and was in many ways as much philosophical as empirical. Ebbinghaus transformed the study of memory from philosophical reflection on one's own memories into a subject of quantitative scientific analysis. He quantified the relationship between repetitions of memorization, accuracy of recollection and durability of memory. Ebbinghaus described the capabilities of human memory in the first concrete terms by counting how many made up words a person can recall, and measuring the period of time during which they can be recalled. He showed that the trope "practice makes perfect" had a basis in fact, demonstrating that one could recall only 6 or 7 items after brief study, and that the ability to recall these items accurately degraded rapidly over the course of minutes, but with repeated study, one could dramatically increase both the number of items and the period over which they could be recalled. Further, by studying the rate at which recollection failed, he observed that memory degraded in two phases, one in the hours immediately after learning, and one that lasted for months. Thus, his work indicated that long-term memory (LTM) was an extension of short-term memory (STM), but that distinctions existed between the two. (*1*) The work of Ivan Pavlov provided seminal insights into the mechanistic aspects of learning and memory. His use of animal models and simple stimuli with easily quantifiable measures showed that techniques to empirically study the biology of the mind could be developed. His studies of reflex also provided a first hint to the physiological basis of learning. In this work, he developed what would

come to be termed classical conditioning, in which a subject can be taught to associate an arbitrary conditioned stimulus (CS), with a biologically significant, unconditioned stimulus (US) such as food or pain, which normally elicits an unconditioned response (UR) such as salivation or assuming a fear posture. Following conditioning, the CS will elicit a conditioned response (CR) as if the subject was presented with the US. Specifically, Pavlov showed that after repeated sessions in which a bell is rung just prior to giving a dog a taste of food, the dog will associate the ringing bell with the taste of food, and as a result, the dog will salivate in response to the ringing bell, even in the absence of food. Using this method, Pavlov was able to directly examine the formation and persistence of a newly created memory of an association of stimuli. Crucially, he found that conditioning was most effective when the CS was presented just prior to the US, thereby becoming predictive of the US. Another form of learning, in which animals are taught to associate a behavior with its consequences, was discovered by Konorski and studied extensively by Thorndike. In this paradigm, behaviors are punished or reinforced, resulting in the increase or decrease in the probability of the subject displaying the conditioned behavior. This form of learning came to be termed operant conditioning. Based on this work, Thorndike developed a theory of learning that included the idea that all animals learn the same way. Cajal, based upon his careful study of neuroanatomy, first proposed a physiological manifestation of memory in the form of synaptic connections between neurons. Later, Hebb would use behavioral studies as well as measurements of the electrical activity of the brain as the basis of what became known as Hebbian theory. A central component of this

theory is that the firing patterns of neurons affect the efficacy of their synapses. He suggested that when one neuron repeatedly participates in triggering the activity of another, physiological changes in one or both cells would occur, thereby enhancing the ability of the first cell to trigger activity in the second cell. This was shown to be true by Lomo and Bliss in studies of the rabbit hippocampus. In this work, they demonstrated that a single electrical stimulation of the perforant pathway to the dentate gyrus caused excitatory post-synaptic potentials (EPSP) in the neurons of the dentate gyrus. However, preceding such stimulation with a high-frequency train of electrical pulses would cause a change such that the post-synaptic cells of the dentate gyrus would produce stronger and longer lasting EPSPs. Further, this change could last 30 minutes to several hours. Lomo and Bliss termed this phenomenon long-term potentiation (LTP).(*2*) This observation strongly supported Hebbian theory, but the complexity of the mammalian brain hindered subsequent progress. Kandel and his colleagues made significant contributions in work on the cellular and molecular biology of the gill withdrawal reflex of the sea hare Aplysia californica. Having identified the gill withdrawal reflex as a behavior that could be easily observed and triggered, they mapped a simple circuit, composed of relatively large, and easily manipulated neurons, governing the reflex. In this circuit, a presynaptic glutaminergic sensory neuron forms direct synaptic connections with the motor neurons governing the gill withdrawal reflex, as well as synaptic connections with both excitatory and inhibitory interneurons that themselves form synapses with the sensory and motor neurons. Using this system, Kandel and his colleagues demonstrated that pairing stimulation of the siphon (the CS) with an

electric shock to the tail (the US) enhances the gill withdrawal reflex (the CR), and that this enhancement is correlated with changes in synaptic strength. A single Pavlovian pairing of US and CS transiently enhances the ability of the presynaptic sensory neuron to trigger an action potential in the postsynaptic motor neuron. This enhancement requires serotonergic modulatory input triggered by the US. 5 or more pairings leads to formation of a more persistent, intermediate-term memory (ITM) of the association, lasting several hours before the response returns to baseline. Previously quiescent synapses can become active, but the number of synaptic connections remains constant.(*3*) Repeated sessions of 5 or more pairings of siphon stimulation and tail shock, with periods of rest between sessions, leads to stabile, long-term memory (LTM) formation, lasting for days or longer. This LTM forming paradigm results in strengthening of existing synapses as well as the formation of new synapses.(*4*) Thus, Kandel and his colleagues demonstrated that aspects of memory described by Ebbinghaus had invertebrate neural correlates akin to those predicted by Cajal and Hebb. They also showed that invertebrate neurobiology could provide insights that would be, at the time, precluded by the complexity of neural tissues in higher organisms and the obstacles to work in mammalian systems such as slower reproduction and ethical concerns. In this vein, Benzer and his colleagues used the power of Drosophila mutants to pioneer work in behavioral genetics, demonstrating that individual genes could govern complex behaviors. Subsequent work showed that many of these genes are broadly conserved, with similar contributions to behavior in other organisms. (*5-7*) The advent and dissemination of powerful biochemistry and molecular biology methods

since the 1970s has permitted study of learning and memory at the most fundamental level. It is now known that major features of learning and memory are evolutionarily conserved at molecular, cellular, and behavioral levels. Mollusks, insects, rodents and humans all possess phases of associative memory commonly categorized as STM, ITM, and LTM. At the cellular and molecular level, major mechanisms involved in the formation of each type of memory are also conserved. The initial, and least stabile phase of memory formation involves modification of proteins already present at existing synapses via the cAMP pathway. Neural activity relevant to learning activates adenylyl cyclase (AC) and phospholipase-C (PLC). AC produces cyclic adenosine mono-phosphate (cAMP), and cAMP in turn activates the cAMP-dependent protein kinase (pKA). pKA phosphorylates K^+ channels, reducing the influx of K^+, and prolonging the action potential. PLC cleaves phosphatidylinositol 1,4,5-bisphosphate (PIP_2) yielding inositol 1, 4, 5-trisphosphate (IP_3) and diacylglycerol (DAG). DAG activates protein kinase C (pKC), leading to the opening of L-type Ca^{2+} channels, which have a prolonged response, and thus lead to greater calcium influx.(*8*) Consolidation of memories into LTM, involves the cAMP signaling pathway as well. During consolidation, pKA recruits the mitogen activated protein (MAP) kinase. Together, pKA and MAP translocate to the nucleus, where they phosphorylate the transcription factor cAMP response promoter element binding protein (CREB). This allows CREB to bind the cAMP response promoter element (CRE), activating transcription of genes under its control. These genes have numerous activities and trigger biochemical pathways that act in concert to stabilize the enhancement of synaptic efficacy. Key amongst these CREB regulated genes is

the transcription factor CCAAT-enhancer-binding protein (C/EBP), which itself activates a cascade that leads to the growth of new synaptic connections. (*6-9*)

Many Molecular and Genetic Tools Are Available For Studies Of Drosophila Learning And Memory

Drosophila was a well-developed model organism for use in genetic studies long before modern molecular biology became widely used. Researchers have been collecting and inbreeding *Drosophila* mutants since the early 1900s, and now have cultured strains carrying mutations in most genes. Furthermore, many methods with which to generate random and directed mutations in flies exist. The mobile DNA P-element found in *Drosophila* has been of particular utility for generating desirable mutants. Random insertion of P-elements within genes, or the excision of previously inserted P-elements within genes can produce mutations. Constructs carrying transgenes flanked by P-element sequences that permit insertion into the genome are a widely used method for delivering transgenes to the *Drosophila* genome. Several methods for tissue specific expression of transgenes in flies have been developed. Gene targeting has been developed in *Drosophila*, but its use remains infrequent. A widely used method for tissue specific expression is the binary GAL4 / UAS system, in which the yeast transcriptional activator GAL4 is placed under the control of a Drosophila regulatory element, and drives expression of a transgene at a separate genomic location under the control of the yeast upstream activation sequence (UAS). A large number of random P element

insertions carrying GAL4 enhancer trap constructs have been collected and cultured. Such insertions will drive expression of GAL4 according to regulatory elements in the genomic context surrounding the insertion. This can result in GAL4 expression closely mimicking the spatial and temporal expression patterns of genes near the enhancer trap insertion.(*10*) Recently, a consortium from the Howard Hughes Medical Institute's Janelia farm have cloned genomic fragments spanning much of the *Drosophila* genome, and placed these fragments immediately up-stream of a core promoter driving GAL4 expression. This promoter will become active only if the cloned fragment contains an enhancer. These constructs are then inserted into the Drosophila genome at a fixed location so as to avoid position effect variegation from random insertion. These constructs will drive GAL4 expression in patterns matching that of genes controlled by elements in the cloned fragment. Using this collection, the consortium is identifying neural expression patterns of genes, and mapping the circuitry of the brain. Much of the husbandry, imaging, and analysis in this work has been automated.(*11*) The binary nature of the GAL4/UAS and similar systems available in *Drosophila* makes such GAL4 collections enormously powerful tools, as experiments involving expression of different transgenes in defined tissues can be accomplished simply by crossing flies expressing GAL4 in the desired tissue to flies harboring the transgenes under UAS control. Similarly, a single transgene can be expressed in different tissues through simple crosses. Binary expression systems have been used to map circuitry relevant to learning and memory using GAL4 expression under the control of genes required for these processes to drive expression of UAS-GFP. Another approach uses

screening for learning or memory defects in flies carrying GAL4 enhancer traps driving expression of transgenes that silence neural activity under UAS control. One such transgene, the dominant negative, temperature sensitive dynamin mutant Shibere^{ts} (Shi^{ts}), which reversibly blocks synaptic transmission when shifted to a temperature above 29°, has been used to map the behavioral contribution of various neural circuits. GAL4 enhancer trap lines from crosses exhibiting defects can then be crossed to flies harboring UAS-GFP or other reporter transgenes, thereby marking the neurons involved. Photoactivatible GFP variants (PA-GFP) permit tracing of individual neurons. Neurons expressing PA-GFP and innervating a neuropil of interest are illuminated such that PA-GFP is converted to its fluorescent state. This activated PA-GFP is then allowed to diffuse to the cell bodies of neurons innervating the neuropil. An individual soma is selected, and PA-GFP converting illumination is maintained at that soma while fluorescence in other neurons decays. In this way, the branching patterns of individual neurons, and wiring diagrams of circuits can be precisely mapped.(*12*) Neural activity in these circuits can be recorded during odor presentation by expressing genetically encoded fluorescent calcium sensors under the control of the same GAL4 lines, and using confocal or two-photon microscopy to collect images of the brain from intact heads as they are presented with odors.(*13*, *14*) Neural activity in the same cells can also be directly controlled optogenetically by crossing the GAL4 line to flies expressing channelrhodopsin-2 or halorhodopsin under UAS control.(*15*) Further, a number of methods for conditional expression of transgenes in *Drosophila* are available. HSP70-GAL4 drives GAL4 in all tissues when flies are subjected to elevated

temperature. GAL80ts blocks transcription of genes under UAS control until flies expressing it are subjected to elevated temperature. Thus, by combining ubiquitously expressed GAL80ts with a GAL4 line of interest, expression of genes under UAS control can be manipulated spatially and temporally. For developmental studies, mosaic analysis with a repressible cell marker (MARCM) can be used to drive expression of transgenes in cells from specific lineages. These examples represent a subset of the genetic tools available in *Drosophila*, but underscore the modular design and combinatorial potential available in flies. This illustrates why *Drosophila* has become a uniquely well-suited model organism for experimental manipulation and study of learning and memory.

Paradigms For Behavioral Studies In *Drosophila*

Drosophila can form tactile, spatial, visual, gustatory, and olfactory memories in operant and/or classical conditioning paradigms. (*7*, *13*, *16*) A variety of behavioral methods has been developed to study learning and memory in *Drosophila*, with a broad range of complexity and scale. To study visual learning, flight simulators have been developed, in which individual tethered flies can learn to associate a shape or color with a noxious stimulus in the form of heat delivered by infrared (IR) laser light.(*17*) Others developed an arena, accommodating one or many larvae or adults, in which areas illuminated with light of a particular color contain sucrose. Large groups of flies can be trained and tested for their memory of the association of color and sucrose using this setup.(*18*) To study memory of place,

an analog of the Morris water maze has been developed in which an array of LEDs surround an arena in which a particular part of the floor is cooled, while the rest of the arena is heated. The ability of flies to find the cool location improves with repeated trials. By manipulating the visual scene presented on the LED array, flies can be tested for their ability to find the cool location when the visual cues are altered or disappear.(*19*) Drosophila can be conditioned to associate a taste with a noxious heat. *Drosophila* respond to presentation of sugars to gustatory sensory neurons in their tarsi by extending the proboscis. This proboscis extension reflex (PER) can be used to evaluate gustatory associative learning. By pairing presentation of a tastant to the tarsi with painfull IR laser light to the antennae, flies can be conditioned to associate the tastant with the noxious heat. Following conditioning, PER is suppressed if the conditioned tastant is mixed into a sugar solution presented to the tarsi.(*20*) Male *Drosophila* readily attempt to mate with appropriate and inappropriate targets, including sexually immature males and females. However, after presentation with a previously mated adult female, the male will exhibit decreased courtship attempts with females, but not immature males for 2-3 hours. Repeated presentation can result in decreased courtship attempts lasting as long as 8 days. This learning is associative, with cuticle hydrocarbons the being the CS, and the male deposited hormone cis-vinyl acetate (cVA) being the best characterized aversive US.(*21*) Recent advances in computational video analysis now permit simultaneous behavioral analysis and/or manipulation of individual flies within large freely behaving groups.(*22*) Exceedingly rich data sets can be rapidly collected with such methods, as they

permit the assessment of group dynamics, and/or high-throughput examination of individual behavior. However, aversive olfactory classical conditioning was one of the first paradigms developed for studies of *Drosophila* learning and memory, and remains one of the most widely used.(*5*) It is also the most developed and best understood paradigm in widespread use. In this paradigm, flies are presented with two odors, one of which is accompanied by inescapable electric shock. Conditioning occurs in the dark to ensure that the odors being presented are the most salient stimuli other than the US. Conditioned flies are then tested for odor choice in the dark, using a T-maze in which the flies begin at the bottom of the T, with one arm containing the conditioned odor, and the other arm containing the control odor or a novel odor. Odor concentrations are identical to those used during conditioning. Innate negative geotaxis leads the flies to the choice point. Following conditioning, the number of flies in the arm containing the conditioned odor is significantly reduced. Learning and memory can be evaluated by counting the number of flies in each arm, and then calculating a performance index (PI). PI is defined as the fraction of flies that avoid the conditioned odor, minus the fraction of flies that avoid the control odor. This means that a PI of 0 represents no learning, and a 50:50 distribution of flies in the two arms, while a PI of 1 represents perfect learning, with all flies avoiding the conditioned odor. This aversive olfactory conditioning method allows many flies to be simultaneously trained and tested, thereby vaulting over one of the major hurdles of behavioral work in other animal models. (*7*, *13*, *23*) Recently, a semi-automated method for performing aversive olfactory conditioning of *Drosophila* has been developed, permitting multiple specimens to be conditioned

simultaneously, and in a reproducible manner.(*24*) This list is not exhaustive, and new paradigms with which to study *Drosophila* behavior continue to be developed. The diversity of methods with which to study and manipulate *Drosophila* sensation, learning, and memory demonstrates the surprising capabilities of the drosophila brain. This diversity also illustrates the exceptional suitability of *Drosophila* as a model organism with which to study learning and memory.

Olfactory Learning and Memory in Drosophila

Quinn and Tully used aversive olfactory conditioning to study the genetics of learning and memory, and were able to identify several genes that, when mutated, caused defects in specific aspects of these processes. (*5*, *6*, *25-27*) Subsequent work has identified many more mutations and P-element insertions that result in learning or memory defects. These have in turn been used to map the neural circuits involved in olfactory learning and memory in *Drosophila*. Identification of the genes and circuits involved has greatly contributed to understanding the molecular mechanisms underlying synaptic plasticity as well as learning and memory more broadly. Many of these molecular mechanisms are evolutionarily conserved, and analogous circuitry exists in many insect and non-insect species.(*6*) Drosophila aversive olfactory memory can be broken down into 4 distinct phases following learning: STM, middle-term memory (MTM), anesthesia resistant memory (ARM), and LTM. *Drosophila* STM is present immediately after a single session of aversive olfactory classical conditioning, and is subject to disruption by anesthesia. STM

persists for 30-60 minutes, and does not require transcription or protein synthesis. A single training session will also produce MTM lasting several hours, but by 24 hours after training, PIs return to baseline. MTM can be disrupted by anesthesia or reversal training, in which previously conditioned flies are exposed to the same odors used during conditioning, but the previously conditioned odor is now neutral and the previously neutral odor is now paired with shock. Following 5 -10 conditioning sessions with no intervening rest period (Massed training), STM, MTM, and a more durable form of memory are produced. This durable form of memory is not subject to disruption, and is therefore termed anesthesia resistant memory (ARM). ARM can last for as long as 4 days before PI diminishes completely, and does not require protein synthesis. 5-10 conditioning sessions with intervening periods of rest (Spaced training) produces STM, ARM, and LTM. LTM is distinguished from ARM by its requirement for protein synthesis, and CREB. LTM can persist for many days and is not subject to disruption by anesthesia. Spaced training also produces better PI immediately following conditioning than massed training, which in turn produces better PI than a single conditioning session. (*6*, *7*, *16*) However despite its crucial role in advancing the study of learning and memory, recent studies using other learning paradigms suggest that some aspects of *Drosophila* memory formation resulting from aversive olfactory classical conditioning, long thought to be generic to all modes of learning, may in fact differ between learning paradigms. For instance, appetitive olfactory conditioning, in which flies are taught to associate an odor with the reward of sugar, can produce LTM after a single training session. Though many such differences were reported long ago, they have only recently

begun to be investigated in detail, and no *Drosophila* training paradigm has been as widely used and well studied as aversive olfactory classical conditioning. (*28*) The automation of many newer learning paradigms may speed the identification of aspects of memory unique to different sensory modalities, unconditioned stimuli, or training methods, but significant work remains to categorize some features of learning and memory in *Drosophila* as truly general or more specific to a given paradigm. Nonetheless, aversive olfactory classical conditioning remains the standard to which other paradigms are compared.

Genetic Analysis of Drosophila Olfactory Memory

There are unique genetic requirements for aversive olfactory classical conditioning learning, and for each form of memory it produces. Indeed, much of what is known about the phases of *Drosophila* learning and memory comes from genetic analysis of mutants exhibiting behavioral defects when subjected to aversive olfactory classical conditioning. Learning itself requires components of the cAMP signaling pathway. Mutants of the drosophila PKA regulatory subunit (PKA-RI), the PKA catalytic subunit (Pka-C1), or the PKC inhibitor (14-3-3ζ) are unable to learn, but are sensitive to shock, and are able to perceive odors. (*6*, *16*, *29*) In addition to cAMP pathway components, fasciclin II (fasII) and latheo(lat) mutants show similar learning defects. Mutants for other genes in the cAMP signaling pathway exhibit both learning and STM defects. The cAMP phosphodiesterase dunce (dnc), and the Ca^{2+} / calmodulin dependent adenylate cyclase rutabaga (rut), are required for

learning and STM, but not MTM. Mutants of the integrin scab (scb) or neurofibromin 1 (NF1) also show learning and STM defects, but retain some MTM. MTM was itself confirmed as a distinct phase of memory based upon analysis of the amnesiac (amn) mutant. STM decays rapidly, and ARM appears gradually around 60 minutes after conditioning, but observations of memory retention showed a smooth decline following training. It had therefore been surmised that another disruptable phase of memory existed, and that the additive contributions of STM, ARM and this surmised phase produced the observed retention curve. Careful study of amn flies shows that they are able to learn, have near normal STM immediately following a single conditioning session, and exhibited memory retention 6-7 hours later. However, amn flies exhibit significantly reduced memory during the intervening period. Further, when reversal training is conducted at various time points following conditioning, the memory of normal flies is shown to be subject to disruption during the same period. Moreover, the retention curves of normal flies subjected to reversal training, and of amn flies are very similar, indicating that a disruptable phase of memory between STM and ARM exists and requires the amn gene product. Mechanistically, amn encodes a neuropeptide that binds a G-protein coupled receptor, thereby activating adenylate cyclase.(*30*) When flies harboring a temperature sensitive Pka-C1 ($DC0^{ts}$) are shifted to the restrictive temperature just before beginning behavioral experiments, they memory defects that are indistinguishable from amn. In this way, Pka-C1 and amn mutants demonstrate that MTM is an early form of memory that persists into the onset of ARM, that is distinct from STM, and that has unique genetic requirements. In many ways, ARM

resembles the other form of consolidated memory, LTM. However, While ARM decays within 4 days, LTM can persist for more than a week. The radish (rsh) gene is the only gene known to be specifically required for ARM induced by aversive olfactory classical conditioning and not for LTM. However, it has been reported that in other learning paradigms, rsh mutants display memory defects as early as 3 minutes after learning, suggesting that rsh may not be required generally for memory consolidation, and that the requirement for rsh in ARM is specific to olfactory aversion.(*31*) This view is counter to the prevailing model based upon results from aversive olfactory classical conditioning that implicate rsh in a serotonin mediated pathway for memory consolidation parallel to cAMP signaling. (*6*, *7*, *32*) In *Drosophila*, as in other species, administration of the protein synthesis inhibitor cycloheximide (CMX) blocks LTM formation. ARM does not require protein synthesis, and is thus unaffected by CMX administration. LTM also uniquely requires the transcription factors dCREB2, Adf1, and Notch. Additionally, flies require APPL expression specifically for LTM, and not for learning or other types of memory. Though APPL null mutants show defects in response to shock, likely due to developmental defects, conditional expression in the mushroom body of an RNAi inducing hairpin directed against APPL beginning 24 hours prior to conditioning significantly reduces LTM, but not STM or ARM. These flies are able to perceive odor and react to shock normally, indicating that APPL exerts its effect on LTM independent from its developmental role.(*33*) In addition to genes specifically required for distinct types of aversive olfactory memory, a large and growing list of genes have been implicated more broadly in various aspects of learning and

memory. Many of these were identified in screens for genes involved in LTM, but may act more broadly in memory formation. For instance, appetitive and aversive olfactory memory require G-protein coupled signaling triggered by the dopamine receptor (Dop1R1). (*28*, *34*, *35*) Dopaminergic signaling is thought to act through rut during memory formation, resulting in CREB mediated transcriptional changes.(*34*) Thus, flies with genetic disruptions at many downstream points in this pathway exhibit learning and memory defects. However, the critical role of Dop1R1 in olfactory memory is demonstrated by the complete ablation of aversive olfactory memory in Dop1R1 mutants, while mutants of other genes required for learning such as rut only exhibit decreased learning. The requirement for cAMP signaling pathway components at various stages of learning and memory illustrates the central role that this pathway plays in regulating the cellular processes involved in learning and memory. It also mirrors mechanisms found to be at work in classical conditioning of *Aplysia*. (*8*, *34*) cAMP signaling is activated during learning, leading to rapid but temporary changes in synaptic efficacy. These initial changes can be stabilized through feedback loops such as CaMKII autophosphorylation and more permanently by altered gene expression resulting from CREB mediated transcription. Additionally, a number of genes involved in RNA localization and translational regulation are involved in memory formation. These include Staufen, Pumilio, Oskar, Fmr1, and components of the microRNA induced silencing complex (miRISC). Defects in RNA localization and translational control are an emerging area of study in several neurodegenerative diseases and cognitive disorders. In addition to a role in regulating gene expression generally, miRNA may play a crucial

role in regulating translation at the synapse of mRNAs involved in synaptic plasticity. Mounting evidence indicates that regulated RNA localization and synaptic translation is crucial in maintaining appropriate synaptic connectivity. This also suggests a mechanism by which plasticity could be modulated at specific synapses within a single neuron, though proof of such a mechanism is currently lacking.

The Neural Circuitry of Drosophila Olfactory Memory

The neural circuitry involved in aversive olfactory classical conditioning has been the subject of extensive study in *Drosophila*, and is therefore well understood. Much of what is known about this circuitry comes from genetic dissection, and recording from and manipulation of neural activity in selected brain regions using methods I have previously described. *Drosophila* olfactory sensation takes place in hair like structures called sensilla present on the third segment of the antennae, and on the maxillary palps. A total of aproximately 1200 olfactory receptor neurons (ORN) innervate these structures. Each ORN expresses 1 of 59 olfactory receptor (DOR) genes and an invariant olfactory coreceptor (Orco). The DOR gene product complexes with Orco to form a functional DOR. All ORNs expressing the same DOR project to the same bilaterally symmetric glomeruli of both antennal lobes (AL). The antennal lobes are composed of 43 distinct glomeruli, each receiving input from ORNs expressing a fixed set of DORs. In this way, an odorant map is produced in the AL by the distinctive projection patterns of ORNs. Specific spatial patterns in the AL are activated by input from ORNs, and these patterns are determined by the unique

set of DORs triggered by a given odorant. (*14*, *36*) In the AL, ORN axons form synapses with inhibitory and excitatory local interneurons (LN) that integrate information between glomeruli. ORNs and LNs form synapses with projection neurons (PN). Dendritic arbors of a given PN are typically confined to a single glomerulus. PNs relay olfactory information integrated in the AL to the calix (CA) of the mushroom body (MB), and to the lateral horn (LH). (*7*, *28*, *34*) The set of PNs activated by a given odorant can be modified through conditioning, but this modification lasts only minutes.(*37*) The mushroom body has long been known to be involved in memory, as many of the genes required for memory formation are highly expressed in it, and disruptions of the MB's development or function result in memory defects. (*6*, *34*) Whereas the MB is involved in behavioral response to learned olfactory information, the LH is thought to be involved in innate responses to olfactory information such as those triggered by pheromones. (*28*, *38*) The MB is composed of several lobes, each innervated by 1 of 3 distinct Kenyon cell (KC) populations. The cell bodies of KCs are immediately adjacent to the CA, into which all KCs extend dendrites. In the CA, KCs receive olfactory input from PNs. While a stereotyped odorant map exists in the AL, odor representation in the MB differs between individuals, with KCs integrating input from PNs from apparently random AL glomeruli.(*34*) In the MB, olfactory information is represented by distinct patterns of sparse activity across all three 3 major classes of KCs. KCs of each major class extend two axon branches into specific MB lobes. α/β KCs extend one axon branch into the α lobe, and another into the β lobe. Similarly, each α'/β' cell extends one axon branch into the α' lobe and another into the β' lobe. γ KCs extend both

branches into the γ lobe. (*28*, *34*) The sparse pattern of odor evoked KC activity is likely the result of two mechanisms. First, in the locust MB, KCs function as coincidence detectors, integrating input from multiple PNs. Examination of the dendritic arbors of *Drosophila* KCs shows that each KC typically has 7 dendritic claws, each receiving input from a single distinct PN, thereby indicating that the wiring for coincidence detection of PN activity is present in *Drosophila* KCs. Second, KCs also receive inhibitory GABA-ergic input from non-spiking neurons termed anterior paired lateral neurons(APL) throughout the MB, including in the CA. It has recently been shown that APL activity reduces odor evoked KC activity, further supporting the notion that GABA-ergic APL activity tunes KC activity.(*39*) In addition to olfactory information, the MB also receives input from separate aversive and appetitive reinforcement pathways. For many years, it was thought that dopamine only functioned in aversive signaling, with octopamine functioning in appetitive signalling. This idea is now known to be the result of an experimental artifact, and it has been shown that both aversive and appetitive input occurs via dopaminergic signaling through Dop1R1, albeit through distinct circuits. Dopaminergic protocerebral paired lateral 1 (PPL1) and protocerebral anteriorl lateral (PAM) neurons project to the α'/β', α/β, and γ lobes of the MB, and provide aversive reinforcement input.(*15*, *40*, *41*) Electric shock of the sort used in aversive olfactory classical conditioning activates these neurons. (*40*, *42*) However, restoring expression of Dop1R1 only in the γ lobe of Dop1R1 mutant flies is sufficient for aversive memory formation.(*43*) Dopaminergic input involved in appetitive reinforcement comes from a set of PAM neurons distinct from those involved in

aversive reinforcement. These appetitive dopaminergic PAM neurons project only to the β, β', and γ lobes of the MB, and neurotransmission from these neurons during odor exposure leads to appetitive memory formation. Appetitive PAM neurons respond to octopamine, and octopaminergic reward signaling occurs through these PAM neurons. (*44*, *45*) In addition to olfactory and reinforcement inputs, the MB also receives input from the dorsal paired medial (DPM) neurons. DPM neurons project throughout the MB, except to the CA. DPM neurons express the amn gene, and blocking neurotransmission in DPM neurons with shi^{ts} does not affect learning, but prevents consolidation of olfactory memories. Although they express the neuropeptide amn, DPM neurons are reported to be serotonergic. According to this report, serotonin signaling via the serotonin receptor (5-HT1A), expressed in α/β cells. is required for ARM.(*32*) Interestingly, DPM neurons form gap junctions with APL neurons, and blocking expression of innexins required for the formation of gap junctions in either APL or DPM neurons results in MTM defects.(*46*) However, blockade of APL neurotransmission using shi^{ts} inhibits STM but not LTM.(*47*) Taken together, data regarding DPM function implicate it in the process of consolidating memories.(*28*) Experiments in which shi^{ts} is used to block neurotransmission in each of the lobes of the MB at specific times during learning, memory consolidation, and behavioral testing, reveal that distinct MB regions are required for various aspects of learning, memory storage, and retrieval. Neurotransmission in α'/β' KCs is required during, and immediately after conditioning. However, blockade of neurotransmission in α'/β' KCs during testing has no effect. It therefore appears that α'/β' KCs are required for acquisition, and the earliest stage of memory

formation, but are not the ultimate locus of memory storage. Blockade of the neurotransmission in α/β and γ cells during aversive olfactory conditioning does not disrupt memory formation, but α/β neurotransmission is required during testing. Expression of rut in the γ lobes of rut null mutants rescues STM, but not LTM defects, whereas expression of rut in α/β and γ lobes rescues STM and LTM. Taken together, these results support a model for aversive olfactory memory storage in which olfactory memory is initially encoded in the α'/β' lobes and quickly transferred to the γ lobes. APL and DPM signaling help maintain the memory trace while it is consolidated, and transferred to the α/β lobes where it is stored as a stabile LTM.(*28*, *34*)

Though the circuitry of the drosophila brain is complex enough to produce a surprising array of behaviors, it is simple enough that one can identify the contributions of small groups of cells to these behaviors. Genetic dissection of olfactory memory has identified biochemical pathways and neural tissues involved in memory formation in drosophila. This work demonstrates that the MBs are the primary center for encoding and storing olfactory memory. The changes in behavior resulting from olfactory conditioning stem from altered patterns of neural activity in circuits residing within the MBs. Such insights demonstrate the unique suitability and capabilities of *Drosophila* as a model system for studying the physiology of learning and memory. Work in flies has shown that modification of existing synaptic proteins can produce the changes in neural activity underlying memory, but only transiently. LTM formation requires transcription and translation, and can

involve changes in synaptic number or patterns of synaptic connectivity. Thus, the importance of understanding the ways in which gene regulation is altered during LTM formation is clear. Again, the uniquely powerful combination of behavioral, genetic, and other tools available in *Drosophila* make it a premier model organism with which to study gene regulation in response to LTM formation.

Part II: Short Noncoding RNAs In *Drosophila* Memory Formation

A Variety Of Short, Non-Protein Coding RNAs Regulate Gene Expression In Animals

All somatic cells within a given organism share a common genome. Yet, an astounding diversity of sizes, shapes, and behaviors exists amongst the various cell types that organism possesses. The meter long motor neurons innervating muscles in the feet of mammals bear little resemblance in form or function to the fibroblasts of the same animal, but fibroblasts can be experimentally coaxed into becoming motor neurons, thereby demonstrating that they share the same genetic potential.(*48*) Moreover, a single cell type can exhibit diverse behaviors. For instance, upon infection, a relatively quiescent immune cell can become mobile, actively stalking pathogenic bacteria, engulfing the bacteria once it is caught, producing enzymes to destroy the bacteria, and secreting factors to recruit other immune cells to the site of infection. Both the diversity of cell types and the array of behaviors of a given cell type are largely achieved through dynamically regulated gene expression. As altered gene expression is required for LTM formation,

understanding the mechanisms by which this is achieved provides deep insight into the physiology of memory and cognitive disorders. Work in recent years has significantly broadened and deepened our knowledge of the regulatory potential of animal cells. Much of this new understanding stems from the continual identification of new and surprising activities of non-protein coding RNAs. The discovery of several families of short non-coding RNAs (sRNA) that regulate gene expression through related but distinct mechanisms is a major development in this vein. sRNAs are known to regulate gene expression via transcriptional gene silencing (TGS), post transcriptional gene silencing (PTGS), through DNA and histone modification, heterochromatin formation, and new roles continue to be discovered. (*49-52*) Three major classes of sRNAs have been identified: short interfering RNAs (siRNAs), micro-RNAs (miRNAs), and piwi-interacting RNAs (piRNAs). Though significant distinctions exist between species, all three classes are broadly conserved across animal phyla. The three major sRNA classes are distinguished by their biogenesis, size ranges, mechanisms of action, and protein complexes with which they associate. siRNAs and miRNAs are derived from double stranded RNA (dsRNA) precursors, while piRNAs are produced from single stranded RNAs transcribed from repetitive elements and transposons, or via a mechanism that amplifies previously produced piRNAs. All three sRNA classes act via direct binding to members of the Argonaute family of effector proteins. Argonaute proteins and their bound sRNAs are incorporated into multiprotein complexes termed RNA induced silencing complexes (RISCs). In *Drosophila*, each sRNA class is bound by distinct sets of Argonaute proteins, and these interactions are the major

determinants of the ultimate regulatory effect of sRNAs.(*52*) The Argonaute family of proteins is divided into the Ago and PIWI clades. *Drosophila* have two members of the Ago clade (Argonaute-1 (Ago1) and Argonaute-2 (Ago2)), and 3 PIWI clade members (P-element induced wimpy testes (piwi), Aubergine (Aub), and Argonaute-3 (Ago3)). The overwhelming majority of miRNAs bind Ago1 and negatively regulate mRNA expression via translational silencing.(*52*) siRNAs bind Ago2 and are involved in a variety of processes including post-transcriptional gene silencing (PTGS), heterochromatin regulation, PolII pausing, viral defense, and genome protection against selfish genetic elements. In Drosophila, both endogenously produced siRNAs (esiRNAs) and exogenously introduced siRNAs direct Ago2 mediated destruction of target RNAs. However, esiRNAs can also guide Ago2 in its other activities just mentioned.(*51*, *53-57*) piRNAs can bind all three members of the PIWI clade and largely function in the germline to silence transposable elements (TEs).(*52*, *56*) However, instances of piRNA activity in somatic cells, both in TE silencing and in gene regulation, have been reported. (*58-62*) The mechanism of sRNA biogenesis and sequence features of a given sRNA direct it into binding with a given Argonaute protein, and thereby largely dictate the ultimate function of the sRNA. (*63-65*) Some sRNAs are expressed at barely detectible levels, while others are highly expressed. Moreover, sRNA expression is dynamic, responding to and directing various processes from development to immunity.(*51*, *66*, *67*) The diversity and abundance of sRNAs adds significant complexity to the task of understanding gene regulation. However, it is now clear that sRNAs compose an ancient and vital component animal physiology. Recent

studies have identified both miRNAs and piRNAs as important factors in *Drosophila* neurobiology.(*59*, *68-70*) The array of known esiRNA functions continues to expand, and recent work suggests that they are involved in olfactory sensation, learning, and other neural processes. (*71-73*) Though some aspects of sRNA biogenesis and function are well understood, these phenomena are relatively recent discoveries, and much work remains. Further, the relevance of individual sRNAs for particular biological processes have been characterized, but analysis of broader sets of sRNAs in most processes remain lacking.

Biogenesis And Function Of siRNAs

The first hints at the regulatory potential of sRNAs came in the 1980s from experiments in plants.(*74*) When two constructs carrying separate transgenes, but sharing regulatory sequences, were introduced into the same plant, expression of both transgenes was inhibited. However, when the transgenes were segregated in progeny, both were expressed. This was shown to be the result of both DNA methylation and PTGS. Both processes involve production of sRNAs that base pair with regulatory sequences and RNA products of the transgenes respectively. (*75-77*) Further, it was shown that transgenic expression of truncated viral coat proteins in plants lead to immunity from the corresponding virus through the destruction of both transgenic and viral RNAs encoding the coat protein.(*78*) These initial observations in plants were followed by the discovery of similar phenomena in virtually every animal examined. Though these processes in plants share features

with similar systems in animals, significant distinctions exist. I mention sRNA mediated silencing in plants here to underscore the ancient origin of these mechanisms, and to provide historical context. Although sRNAs remain the focus of active and vibrant work in plants, critical early advances in understanding sRNA mediated gene silencing came from *C. elegans* and *Drosophila,* and these species remain premier model organisms for studies of these phenomena. I will therefore restrict my subsequent discussion of sRNAs to animals unless otherwise explicitly stated. Early work in plants and animals indicated that genes could be silenced in an RNA dependent manner, but the relative ability of sense and antisense single stranded RNA (ssRNA), and double stranded RNA (dsRNA) was not yet clear.(*79*) In work designed to explore this issue, Fire et al showed that neither strand of ssRNA induced robust silencing when injected into *C. elegans*, but potent and sequence specific silencing was induced by introduction of dsRNA homologous to target RNA sequences.(*80*) This mechanism is known as RNA interference (RNAi). The potential utility of a means with which to experimentally silence genes of choice was clear, and thus others worked to duplicate this result in other animals. This work has shown the mechanism to be conserved from nematodes to humans. (*49*, *52*, *64*) Significant progress toward understanding the mechanism of RNAi soon came from work in *Drosophila*. Several papers published in rapid succession showed that introduction of perfectly matched dsRNA into *Drosophila* cell extracts leads to processing of the dsRNA into 21-23nt fragments from both strands, even in the absence of target mRNA. These short RNA fragments are termed short interfering RNA (siRNA). Further, cleavage of mRNAs matching the dsRNA in sequence occurs

only within the region corresponding to the dsRNA sequence, and at intervals of 21-23nt. This observation reflects the role siRNAs play as guides, base pairing with targeted RNAs, and thereby directing silencing machinery to the appropriate target. (*81-84*) This silencing machinery consists of the siRNA and a multi-protein complex termed the RNA induced silencing complex (RISC). siRNAs are produced from long dsRNA precursors via RNAse III type processing by a Dicer protein. In Drosophila, two Dicer genes are present, Dicer-1 (Dcr1) and Dicer-2 (Dcr2). Cleavage of such long dsRNA with extensive complementarity by Dcr2 yields duplexes of 21-23nt with 2nt 3' OH overhangs, and 5' monophosphates. However, the vast majority of Dcr2 cleaved duplexes are composed of 21nt ssRNAs.(*51*) Following cleavage by Dcr2, the duplex is loaded into the effector protein Argonaute-2 (Ago2). One strand of the dsRNA, dubbed the passenger strand, is discarded and degraded. Ago2 retains the other strand such that the siRNA is able to base pair with target RNAs, thereby guiding Ago2 to the target.(*52*) The Ago2 retained strand is therefore termed the guide strand. The strand with a less stably base paired 5' end is preferentially retained as the guide strand. The details of siRNA excision from precursor RNAs, and subsequent loading into Argonaute proteins differ between species, though common themes exist. This phenomenon is perhaps best understood in flies. In *Drosophila*, the R2D2 protein functions in sensing the relative stability of the 5' ends of the two dsRNA strands. R2D2 thus plays an important role in strand selection and loading of Ago2.(*52*) R2D2 may function in localizing Dcr2, and R2D2 stability is Dcr2 dependent.(*85*) The D isoform of another dsRNA binding protein, Loquacious (loqs-PD) is also required for siRNA production via Dcr2. Its

function is less clear, but it is believed that loqs-PD facilitates Dcr2 binding of dsRNAs, and that this step is downstream of R2D2's activity.(*86*) The heat shock cognate protein 70kDa-heat shock protein 90 (Hsc70-Hsp90) chaperone complex holds Ago2 in a conformation able to accept the duplex. This Hsc70-Hsp90 mediated conformational change of Ago2 is ATP dependent, and results in the transfer of the Dcr2 cleaved duplex to Ago2. The passenger strand is subsequently cleaved and separated from the guide strand within Ago2 in an ATP independent manner. (*87*, *88*) Following passenger strand cleavage by Ago2, a complex of the Translin and Trax proteins termed the Component 3 Promoter of RISC (C3PO) removes the passenger strand from Ago2. The interaction of C3PO and siRNA loaded Ago2 yields an active siRNA containing RISC (siRISC) that is able to cleave target mRNA.(*89*) The final step of in vivo maturation of siRISCs is 2' O-methylation of the 3' nucleotide of Ago2 bound siRNAs by the methyl transferase DmHen1.(*90*, *91*) Uridines are added to the 3' ends of sRNAs with extensive complementarity to their targets by terminal urydil transferases, and thereby marked for destruction. 2' O-methylation of the 3' nucleotide protects Ago2 bound siRNAs from nucleotide addition and exonuclealytic shortening, and thus stabilizes siRNAs.(*91*) Ago2 cleaves target RNAs with perfect, or nearly perfect complementarity to the siRNA with which it is loaded via an endonuclease activity, leading to the destruction of the target RNA. siRNA directed cleavage of mRNAs by Ago2 is a multi-turnover process, thus siRNAs are potent post-transcriptional silencers of gene expression.(*52*) As such, increased siRNA levels resulting from stabilization by 2' O-methylation of the 3' nucleotide enhances silencing. During the early 2000's siRNA induced RNAi was

developed into one of the most important and widely used methods for manipulating gene expression in animal cells. Researchers are now able to experimentally inactivate any mRNA for which they have sequence information by transfecting corresponding siRNAs, longer dsRNAs, or by generating animals that harbor transgenes producing dsRNA matching the target mRNA sequence. As this system must have evolved with a natural function involving exogenous RNAs in *Drosophila*, the siRNA pathway was long thought to function in viral defense, and indeed it does.(*92*, *93*) However, the siRNA pathway does not act solely via exogenous RNA. With the spread of massively parallel DNA sequencing technology in the late 2000's and its use in sequencing sRNAs, researchers identified endogenous sRNAs bound to Ago2 that matched genomic sequences in *Drosophila.* Equivalent endogenous sRNAs were also found in mice, worms, humans, and other animals. (*50*, *94-97*) As these sRNAs bind the siRNA effector protein Ago2 in *Drosophila*, and share the 21-22nt size range of siRNAs, these sRNAs are termed endogenous siRNAs (esiRNA). esiRNAs arise from a number of sources, and appear to have functions in addition to guiding Ago2 mediated destruction of target mRNAs. As is the case with siRNAs, esiRNAs are produced from long dsRNAs exhibiting extensive complementarity between strands. Such dsRNAs can be generated by annealing of RNAs transcribed from distant genomic locations, or from bi-directional transcription at loci where genes, pseudogenes, or ncRNAs reside on both strands. Long inverted repeats can form hairpin secondary structures with extensive complementarity and also produce esiRNAs. Lastly, heterochromatin, transposons, and other repetitive elements generate and are targeted by

esiRNAs.(*94*, *96*, *97*) esiRNAs can act in PTGS much as exogenous siRNAs. However, such instances appear to be rare. Czech et al reported that esiRNAs generated by a locus, dubbed esi-2, harboring 20 palindromic repeats of ~260nt produces abundant esiRNAs. Some of the most abundant esi-2 esiRNAs have extensive, though not perfect complementarity with a sequence in the coding region of the mus308 gene. Analysis of mus308 cleavage products indicated that they correspond to siRNA type cleavage directed by esi-2 esiRNAs. Further, Ago2 and Dcr2 mutants both exhibited elevated mus308 expression.(*94*) However, despite the large number of esiRNAs cloned thus far, this remains the only well documented case in which esiRNAs regulate endogenous gene expression through PTGS in *Drosophila*.(*98*) The esiRNA pathway also appears regulate gene expression transcriptionally through Pol II pausing. The negative elongating factor E (Nelf-E), causes transcriptional pausing via interactions with Pol II. Nelf-E co-immunoprecipitates with Ago2 and Dcr2 as well as Pol II. Nelf-E and Ago2 association with Pol II is disrupted in cells depleted of Dcr2. Depletion of Dcr2 disrupts transcriptional control of genes that have paused Pol II. Functioning Dcr2 and Ago2 are required for proper transcriptional silencing of heat shock loci under normal conditions. Single amino acid mutations that disrupt the helicase or dicing activities of Dcr2, or the slicing activity of Ago2, result in increased levels of heat shock protein transcripts under normal conditions. The role of esiRNAs themselves in Pol II pausing remains unclear. However, the presence of Ago2 bound antisense esiRNAs mapping to heat shock promoters suggests that they are involved in this process.(*54*) However, the major function of *Drosophila* esiRNAs appears to be in

silencing selfish genetic elements in somatic cells. The bulk of Ago2 bound esiRNAs correspond to transposon sequences. Loss of Ago2 or Dcr2 corresponds with elevated transcript levels of some transposons. (*94*, *95*, *99*, *100*) Heterochromatin is largely composed of transposons and repetitive elements in *Drosophila*. Blocking biogenesis or function of esiRNAs with viral proteins alters patterns of H3K9 methylation and disrupts heterochromatin formation. These results are also seen in Dcr2, R2D2, and Ago2 mutants.(*53*) Though exogenously introduced siRNAs have been a widely used experimental tool for over a decade, much remains to be understood about the origin and functions of esiRNAs. Further, as components of the siRNA mediated silencing pathway are known to interact with many other proteins, both within and outside of RISCs, it is not surprising to find them implicated in a variety of biological processes. Recent work has shown that Ago2 unexpectedly functions in processes from alternative splicing to chromosomal looping. However, it appears that esiRNAs and the slicer function of Ago2 are dispensable for these activities.(*55*, *101*) Further study is required to clearly delineate which processes are siRNA dependent, and in which RISC components merely serve as structural components of protein complexes unrelated to their silencing function.

Biogenesis And Function Of miRNAs

Mature miRNAs are ~21-24nt single stranded RNAs, derived from a variety of longer precursors that form duplexes with themselves via hairpin secondary

structures. miRNAs bind Ago family Argonaute proteins, and are thereby incorporated into protein complexes dubbed microRNA induced silencing complexes (miRISC). Animal miRNAs base pair imperfectly with sequences in mRNAs, thereby targeting these mRNAs for post-transcriptional silencing via miRISC activity. Base pairing with the target in the region of miRNA nucleotides 2-8 is sufficient to induce silencing, and the sequence of this region is therefore the major determinant of miRNA target specificity. This region of miRNAs is termed the seed sequence. Sequences outside of the seed seem to play little role in determining which mRNAs a given animal miRNA targets. While *Drosophila* have only 258 miRNA genes, the brevity of the seed sequence ensures that each miRNA targets many mRNAs. Indeed, most animals have a small number of miRNA genes, yet half or more of mRNAs in most animals are subject to silencing by miRNAs. Plant miRNAs usually exhibit extensive complementarity with their targets, and thereby guide post-transcriptional silencing via endonucleolytic cleavage and destruction akin to the activity of siRNAs. Though plant and animals both possess miRNAs, major distinctions exist in miRNA silencing mechanisms between the kingdoms. I will therefore restrict my discussion to animal miRNAs. miRNAs were first identified in C. elegans, but were subsequently found in plants, fungi, and all metazoans. Indeed, many miRNAs are extensively conserved in animals, as are many of the regulatory parings of miRNAs and target genes. The conservation of sequence and functions in developmental gene regulation exhibited by many miRNAs is indicative of their ancient and vital role in controlling gene expression. (*102-105*) For this reason, miRNAs have been a major topic of study since their

discovery. Several mechanisms for miRNA biogenesis have been identified in *Drosophila*. Canonically, miRNAs are transcribed by RNA Polymerase II (PolII) from non-protein coding genes, yielding a hairpin primary miRNA precursor (pri-miRNA). In many cases, multiple pri-miRNAs exist within a single such transcript. In the nucleus, a heterodimer of the Pasha and Drosha proteins, termed the microprocessor, recognizes such hairpin structures and excises a shorter, ~50-70nt miRNA precursor from the pri-miRNA. This shorter precursor is known as a pre-miRNA. Cleavage by Drosha/Pasha results in a ~2nt 3' overhang on the pre-miRNA hairpin, allowing it to be recognized and actively exported to the cytoplasm via Exportin-5 (Exp-5) and Ran-GTP. In the cytoplasm, the RNAse III protein Dicer-1 (Dcr1) cleaves the pre-miRNA ~2 helical turns into the hairpin. This second cleavage event yields a ~22nt dsRNA with 2nt 3' overhangs. (*52*, *63*, *106*, *107*) Specific isoforms of the Loquacious (loqs) protein bind Dcr1, positioning the pre-miRNA properly in Dcr1, and facilitating cleavage. Loss of loqs-PA and loqs-PB is embryonic lethal in *Drosophila*, but loss of loqs-PD is not. loqs-PA or loqs-PB transgene expression can rescue lethality in homozygotic loqs mutants. However, distinct sets of miRNAs are preferentially produced by loqs-PA or loqs-PB transgene expression in homozygotic loqs mutants. Further, the Dcr1 cleavage sites of some miRNAs differ when homozygotic loqs mutants are rescued with loqs-PA vs. loqs-PB transgenes. The distinct sets of Dcr1 cleavage products preferentially generated by loqs-PA and loqs-PB suggest a model in which relative abundance of different miRNAs, or Dcr1 cleavage products from the same pre-miRNA can be modulated through loqs splice site selection.(*108*) Such miRNA variants from the same pre-

miRNA are termed isomiRs. Following cleavage by Dcr1, the duplex is loaded into the effector protein Ago1. As is the case for Ago2, the relative stability of the 5' ends of each strand of the duplex is a major determinant of which strand will be loaded into Ago1 as the guide strand.(*63*, *107*) Ago1 bound miRNAs usually exhibit a 5' uridine, whereas siRNAs and miRNA strands loaded onto Ago2 prefer a 5' cytidine. The strand not retained as the Ago1 guide is termed the miRNA* strand. Unlike Ago2, Ago1 does not cleave the miRNA* strand. However, after it is discarded, the miRNA* strand is still degraded. Nonetheless, a significant fraction of miRNA* species are incorporated into Ago2. This phenomena reveals the role that features other than 5' end stability play in determining strand selection in Ago1 loading. As pre-miRNAs feature bulges at positions of mismatched nucleotides, the dsRNA produced by Dcr1 cleavage will present different structures to Ago1, depending on in which orientation the dsRNA encounters Ago1. Ago1 is preferentially loaded with duplexes featuring an unpaired 5' end, and unpaired bases at nucleotides 8-11. Duplexes presenting perfect matches at nucleotides 8-11 are preferentially loaded onto Ago2(*109*) In this way, the strands of sRNA duplexes are sorted into different RISCs depending on the base pairing between precursor strands. This sorting is further refined by the absence of 2' O-methylation of the 3' nucleotide of Ago1 bound sRNAs. While this modification protects Ago2 bound sRNAs from nucleotide addition, trimming, and destabilization, Ago1 bound sRNAs are not modified, and remain vulnerable to destabilization by this mechanism. Ago1 bound sRNAs exhibiting extensive complementarity with their targets are subjected to 3' nucleotide addition and trimming. In flies lacking the Hen1, the enzyme responsible

for 2' O-methylation of the 3' nucleotide of sRNAs, Ago2 bound sRNAs with extensive complementarity to their targets are also tailed, trimmed, and destabilized. Similarly, artificial introduction of RNA that is perfectly complementary to a given miRNA leads to trimming, tailing, and reduced levels of that miRNA.(*110*) However, regardless of the methylation status of the 3' nucleotide of sRNAs, sRNA – target duplexes without perfectly complementary are bulged at positions of mismatched nucleotides. Such bulging prevents target dependent 3' nucleotide addition and trimming of sRNAs. In this way, target interactions help to purify the sRNA complement of Ago1 and Ago2, ensuring that Ago1 selectively retains only miRNA strands that have bulged base pairing with their targets.(*91*, *109*, *110*)

While pri-miRNA transcripts are the major source of miRNAs, other RNA precursors that form hairpin secondary structures also produce miRNAs. Deep sequencing of sRNAs typically yields large numbers of reads mapping to snoRNAs, rRNAs, and tRNAs. While few examples of regulatory functions for these ncRNA derived reads have been found, some snoRNAs appear to produce genuine miRNAs. snoRNAs form secondary structures and primarily act as antisense guides for enzymes that chemically modify rRNAs, tRNAs, and snRNAs. Recently, analysis of deep sequencing data from plants and evolutionarily distant metazoan species has shown that a conserved mechanism generates sRNAs with distinct and characteristic size ranges from snoRNAs. Some of these snoRNA derived sRNAs (sdRNAs) enter the miRNA pathway and produce miRNA-like sdRNAs that repress translation of seed matching target mRNAs. This mechanism does not require

drosha, but does require Dicer activity. Indeed, miRNA-like sdRNAs are found in metazoans lacking the microprocessor, suggesting that they may represent an ancient and conserved RNAi mechanism. (*111*, *112*) Deep sequencing shows that miRNA-like sdRNAs bind Ago1 in drosophila.(*113*, *114*) As miRNA-like sdRNAs are functional in other metazoans, and so deeply conserved, they are presumed to be functional in *Drosophila* as well.(*111*) Protein coding gene transcripts can also produce miRNAs. In 2007, deep sequencing of *Drosophila* sRNAs revealed miRNA like sRNAs mapping to the splice acceptor and donor sites of certain introns. These sRNAs were thus dubbed mirtrons. Careful analysis revealed that following splicing, debranched introns can form secondary structures that are exported from the nucleus via exportin-5 and cleaved by Dcr1. Unlike canonical miRNA biogenesis, export and Dcr1 cleavage of mirtrons do not require drosha, as the hairpin structure formed by the debranched intron is innately compatible with with these processes. (*115*, *116*) Biogenesis of some mirtrons requires 3' trimming by the exosome. In these cases, the debranched intron base pairs with itself such that a long 3' tail extends beyond the hairpin. The exosome then trims this tail back to the hairpin, yielding the pre-miRNA.(*117*) However, regardless of exosome involvement, mirtrons are loaded into Ago1, and regulate target mRNAs like miRNAs in flies.(*115-117*). Mirtrons were subsequently found in worms, birds, and mammals, suggesting that they are evolutionarily ancient.(*118-120*) As imperfectly base paired hairpin RNA structures abound in animal cells, but only a fraction of these hairpins yield miRNAs, a mechanism must exist for funneling certain hairpins and not others into the miRNA pathway. A recent report indicates that several sequence features within

and downstream of the pri-miRNA hairpin contribute to entry of a hairpin into the microprocessor in humans. However, no single sequence feature or combination of sequence features identified in this study ensured that an arbitrary hairpin would be processed into a miRNA. It therefore appears that the transcriptional origin, sequence features, and structure all contribute to routing hairpin RNAs into the miRNA pathway.(*121*) The rapid increase in acquisition and availability of sRNA sequencing data continually reveals examples of novel mechanisms for miRNA biogenesis. It is therefore unlikely that the mechanisms described here represent the totality of ways that miRNAs are generated, though they are likely to be the most important in *Drosophila*.

miRNA mediated repression of mRNAs is achieved through several mechanisms in Drosophila. Unlike Ago2, which binds siRNAs and has a strong slicer activity, Ago1 drives miRNA mediated silencing by recruiting other proteins that inhibit translation, destabilize mRNAs, or both. The repertoire of proteins comprising and associating with miRISCs is non-uniform and dynamic. Recent work demonstrates that the protein complement of *Drosophila* miRISCs changes in response to extracellular signaling, and that these changes can modulate miRNA mediated silencing of target mRNAs.(*66*) Following loading of the mature miRNA into Ago1 by the miRISC loading complex, which contains the miRNA, Ago1, Dcr1 and loqs, at least two types of active miRISCs can form. These miRISC types are defined by the proteins that associate with Ago1 within them. In the first type, following loading, Ago1 releases Dcr1 and loqs, and associates with GW182. As it contains GW182, a mature miRNA, and Ago1, this complex is known as G-miRISC. In

the second type, following loading of the mature miRNA into Ago1, Dcr1 disassociates from Ago1, but loqs is retained, yielding a miRISC composed of the mature miRNA, Ago1, and loqs. As loqs containing miRISCs sediment with polysomes, such miRISCs are termed P-miRISCs. GW182 and loqs binding to Ago1 are mutually exclusive. Further, G-miRISCs and P-miRISCs interact with distinct sets of proteins, and thereby silence target mRNAs in mechanistically different ways. In G-miRISCs, GW182 localizes the miRISC and its bound target mRNA to cytoplasmic foci containing mRNAs and proteins called P-bodies. P bodies contain enzymes that degrade mRNAs, and are regions in which mRNA silencing occurs. GW182 interacts with poly-A binding protein (PABP), and recruits the CCR4:NOT deadenylase and DCP1:DCP2 decapping enzymes to the mRNA-miRISC complex. Shortening of the poly-A tail by CCR4:NOT leads to decapping, and subsequent 5' to 3' degradation of the mRNA.(*122*, *123*) GW182 can be recruited to mRNAs by proteins other than Ago1, and GW182 recruitment to mRNAs leads to silencing and degradation, even in the absence of Ago1. GW182 thus links the mRNA degradation machinery to miRNA mediated Ago1 binding of target mRNAs.(*66*, *122*) However, GW182 can also induce translational silencing independent from its role in mRNA degradation. Teathering of GW182 to mRNAs prevents 48s and 80s ribosome formation, and therefore inhibits initiation of translation.(*124*, *125*) However, miRNA silencing can occur without GW182. P-miRISCs contain a mature miRNA, Ago1 and loqs-PB, but not GW182, and silence targeted mRNAs by inhibiting translation elongation.(*124*) P-miRISCs associate with polysomes and other dense, non-translating mRNA-protein complexes. While G-miRISC silencing leads to target mRNA destruction, P-miRISC

induced silencing leaves ribosomes poised on the target mRNA, is reversible, and is thought to be involved in keeping certain mRNAs transnationally silent during intracellular transport. (*66*, *126*) In S2 cells, the relative abundance of the two active miRISC types is altered by lipid and PKC signaling. Thus, the regulatory consequence of miRNA mediated silencing is responsive to extracellular cues. Additional study is required to fully understand the mechanisms involved in miRNA mediated silencing, and how these mechanisms are themselves regulated. miRNAs are unquestionably vital players in animal gene regulation, and resolving such questions is therefore a major priority in many disciplines within the life sciences. (*66*, *124*, *127*)

Biogenesis And Function Of piRNAs

piRNAs are 24-32nt RNAs that act as guides for piwi clade argonaute proteins exclusively, and primarily function in transcriptionally silencing repetitive elements and transposons through heterochromatin formation and maintenance. However, piRNAs can also target protein coding genes for silencing, and trigger PTGS. As is the case with other sRNA pathways, piRNA mediated silencing is a deeply conserved mechanism. Mobilization of transposons in the germ line poses a major threat to genomic integrity, and correspondingly, piRNAs are highly expressed in *Drosophila* ovaries and testis. The ovaries have therefore been the primary *Drosophila* tissue in which piRNAs are studied.(*128*, *129*) Once transcribed, the gypsy family of transposons can reinsert themselves into the genome.

Moreover, gypsy family transposons can form particles capable of infecting other cells. Gypsy transposons are thus targeted for piRNA mediated silencing in both oocytes and the somatic ovarian follicle cells, which surround oocytes and support their development. Loss of the piRNA pathway in *Drosophila* ovaries or testis leads to elevated transposon expression and subsequent sterility. Two distinct mechanisms for piRNA biogenesis exist. Primary piRNA biogenesis involves only piwi, and not Aub or Ago3, and is the only mechanism used in follicle cells. In germ cells, an additional mechanism involving Aub and Ago3 amplifies piRNAs. Mounting evidence indicates that both mechanisms are active in somatic cells outside of reproductive tissues. In *Drosophila*, most piRNAs arise from specific loci termed piRNA clusters. piRNA clusters are repositories of transposons and transposon fragments. Deletion of these clusters results in derepression of transposons throughout the genome. Introduction of novel transposons can lead to incorporation of their sequences into piRNA clusters. However, piRNA clusters also exist in widely dispersed euchromatic transposons, and in sequences within the 3'UTRs of some protein coding genes. In oocytes and sperm cells, most clusters are transcribed, frequently bidirectionaly. In follicle cells, the *flamenco* piRNA cluster, and the 3'UTR of the traffic jam (tj) gene are the major source of piRNAs. The *flamenco* cluster contains sequences from long terminal repeat (LTR) transposons, including the gypsy family. The cluster is transcribed via PolII from a single promoter, and generates a long single stranded piRNA precursor RNA that is antisense to the transposon sequences within the cluster. (*61*, *130*) The other major piRNA cluster in follicle cells resides within the 3' UTR of tj.(*128*, *129*) Interestingly,

tj is a transcription factor that drives piwi expression. tj derived piRNAs have a high degree of complementarity to FasIII transcripts, and FasIII is ectopically expressed in tj or piwi mutants. Thus, the tj gene and its piRNA cluster are intimately involved in piRNA mediated silencing.(*131*) A variety of other genes also produce piRNAs from their 3'UTRs, but the regulatory significance of these has not been well studied.(*132*) piRNA precursors are exported to the cytoplasm, where they undergo primary processing by a mechanism that is not yet fully understood. However, the available evidence supports a model for primary piRNA biogenesis in which the 5' end of piRNAs is defined via cleavage by a mitochondrial membrane associated endonuclease named Zucchini (Zuc). Structural and genetic studies of Zuc also support this view. In vitro cleavage of RNAs by Zuc also leaves a 5' monophosphate, which is a feature of piRNAs. (*128*, *129*, *133*, *134*) In somatic follicle cells, 5' cleaved primary piRNA intermediates then enter perinuclear cytoplasmic loci of piRNA processing termed Yb bodies. Yb bodies contain several proteins which are required for piRNA biogenesis and transposon silencing. These proteins include the TUDOR domain containing proteins Yb and Vreteno (vret), the helicase Armitage (Armi), the coshaperone shut down (shu), and piwi. In the Yb body, piwi preferentially binds 5' cleaved piRNA intermediates bearing a 5' Uridine.(*135*) At this stage, the 3' end of the primary piRNA intermediate extends beyond the 32nt' upper limit of typical piRNA length. The 3' end of the primary piRNA intermediate is then trimmed to the 24-32nt size range typical of piRNAs by an unknown exonuclease. piwi interaction with the precursor protects bases within the piwi footprint from removal by this nuclease. 3' trimming is also coupled to 3' O-

methylation of the piRNA by Hen1. The mature piRNA loaded piwi is then imported to the nucleus, where it triggers transcriptional silencing guided by the piRNA. In the germ line and in somatic cells of non-reproductive tissues, piRNA biogenesis differs in a number of ways. First, the flamingo locus does not dominate piRNA production. In germ cells and somatic cells of non-reproductive tissues, all piRNA clusters are active, and guide silencing of an expanded set of targets beyond those present in flamingo. Unlike flamingo, these clusters are often transcribed bidirectionaly. Importantly, in these cell types, Aub and Ago3 act together to amplify piRNAs targeting active transposons through a mechanism termed the ping-pong cycle. The ping-pong amplification loop is initiated by loading of Aub with primary piRNAs that are antisense to active transposons, either produced within the cell, or already bound to Aub and maternally deposited into oocytes. (*136-138*) Such primary piRNAs base pair with transcripts from active transposons and guide cleavage by Aub via its slicer activity. This cleavage yields a new piRNA whose 5' end is defined by the cleavage site, and that is loaded onto Ago3. As primary piRNAs have a uridine bias at the 5' nucleotide, and transposon cleavage by Aub occurs 10nt downstream from the uridine, Ago3 bound piRNAs have an adenosine bias at position 10, and are in the sense orientation. As is the case in primary piRNA biogenesis, the 3' end of the new Ago3 loaded piRNA precursor extends beyond 32nt at this stage. In the current model for ping-pong piRNA amplification, an unknown exonuclease then trims the 3' end of the Ago3 bound precursor to the binding footprint of Ago3. As in primary biogenesis, trimming and 3' 2'O-methylation by Hen1 are coupled. Ago3 can then cleave antisense piRNA cluster transcripts,

thereby defining the 5' end of a new piRNA precursor destined for Aub. Once loaded onto Aub, this new antisense piRNA precursor is trimmed and 3' 2'O-methylated as in the case of Ago3, thereby yielding a mature piRNA and completing the ping-pong loop. (*129*, *137*, *139*, *140*) Unlike follicle cells, germ cells and somatic cells outside of reproductive tissues do not express Yb, and therefore do not have Yb bodies. Instead, piRNA processing in these cells takes place in the nuage, a perinuclear locus related to Yb bodies. However, Handler et al show that the Yb related proteins Brother of Yb (BoYb) and Sister of Yb (SoYb) are required for primary piRNA biogenesis in the germ line, indicating that this family of TUDOR domain containing proteins plays a critical role in piRNA biogenesis.(*141*) The nuage contains many of the same proteins as Yb bodies, including vret, Armi, and shu. The nuage also contains Ago3 and Aub, and loading of these proteins with piRNAs is thought to occur within it. Transposon silencing via ping-pong derived piRNAs obviously occurs posttranscriptionally, as cleavage of transposon transcripts is required for their biogenesis. However, piwi mediated silencing occurs transcriptionally. Piwi's slicer activity is dispensable for silencing by piRNAs, and nuclear localization is required for silencing by piwi. (*131*, *142-144*) In the nucleus, piRNAs act as guides, directing piwi mediated silencing by base pairing with targets. Genomic context determines the mechanism by which piwi acts. PolII transcription occurs in euchromatin, and piwi bound piRNAs base pair with nascent transposon transcripts in this context. In heterochromatin, which is transcriptionally silent, piwi bound piRNAs directly base pair with single stranded DNA. In either case, piwi recruits factors that lead to heterochromatin formation and maintenance. The histone

methyltransferase Su(var)3-9 is recruited by piwi, either directly, or through recruitment of Hp1a, and methylates H3K9.(*145-147*) H3K9 methylation is a heterochromatin mark, and blocks PolII transcription. Maelstrom (mael) influences H3K9 methylation spreading from marks established by piRNA targeting. H3K9 methylation spreading in mael mutants is not however associated with silencing. Transposons in such regions are transcribed, leading to the conclusion that mael acts downstream of H3K9 methylation to silence transcription. Further, genes nearby piRNA targets can be silecned, though they themselves are not subjected to H3K9 methylation, demonstrating the existence of factors other than H3K9 methylation in piRNA mediated silencing.(*145*) In addition to its role in transcriptional silencing, piwi may also act via PTGS. However, it is difficult to distinguish a direct role for piwi in PTGS from its role in piRNA biogenesis, and subsequent Aub and Ago3 mediated PTGS.(*128*) At a cellular level, the piRNA pathway has been implicated in canalization, a process by which phenotypic traits are maintained despite genetic or environmental differences. In this role, it appears that piwi mediated silencing of transposons helps suppress emergence of new genetic variation, and piwi mediated heterochromatin maintenance helps suppress expression of cryptic genotypes. (*148*, *149*) Additionally, growing evidence indicates that the piRNA pathway plays a critical role during development and in maintaining proper cellular differentiation. As was previously mentioned, maternally deposited piRNAs program silencing mechanisms in oocytes. In this case, maternal piRNAs targeting a protein coding gene convert the gene to a piRNA producing locus. The gene continues to be transcribed, but its transcripts are used

as substrates for the ping-pong cycle, and the transcript derived piRNAs can thereby spread silencing to other loci with sequences similar to those of the transcript.(*138*) Thus, in addition to their role in genome defense, piRNAs are a mechanism by which the genome can be epigenetically programmed for silencing.(*150*) Further, knockdown of piwi leads to failure of germ line stem cell maintainance. (*142*, *151*) Ectopic expression of piwi in somatic cells exacerbates tumor growth and leads to acquisition of germ line stem cell traits. (*62*, *152*) Taken together, these studies indicate that the piRNA pathway drives cellular programs that induce or maintain a less differentiated developmental state, and encourage proliferation.

miRNAs In Neurophysiology And Behavior

Insect and human brains differ by orders of magnitude in complexity and cognitive capacity. Yet the difference in gene number between the two is much less substantial. This demonstrates the essential role gene regulation plays in achieving structural and functional complexity of the brain. The nervous system of an animal is composed of cell types that in some cases differ from each other subtly, and in others dramatically. Moreover, within a given cell type, considerable diversity of behavior and responsiveness to external signals exists. Glia and neurons must develop and function together. Neurons must wire properly, and must maintain largely stable wiring diagrams, yet they must also retain the capacity to strengthen or establish new connections, and to weaken or eliminate others. Further, neurons must send and respond to signaling appropriately. Gene regulation lies at the heart

of all of these processes. Though much is understood about gene regulation generally, neurons have unique requirements. Some genes are expressed only within neurons. Interestingly, RNA binding proteins, such as the Drosophila embryonic lethal abnormal vision (ELAV) gene product, comprise the majority of known neuron specific genes. Many genes undergo brain specific splicing, and certain transcripts are localized to particular regions within neurons.(*153*) Recent work in many organisms has shown that sRNAs play central roles in many such neuron specific regulatory mechanisms. The protein products of several genes that are causative in diseases of the nervous system, such as fragile X mental retardation protein (Fmr1) and ataxin-2 (Atx-2), interact with RISC components and are involved in miRNA mediated silencing in neurons.(*68*) miRNAs were originally identified in the context of their temporal regulation of developmental processes. It therefore comes as no surprise that miRNAs regulate many aspects of neurodevelopment in *Drosophila.* miR-8 negatively regulates neuroepithelial expansion and neuroblast transition via control of the epithelial growth factor receptor (EGFR) signaling pathway in glia of the optic lobe. miR-8 silences spitz (spi) via target sites in its 3'UTR. Spitz is a transforming growth factor α (TGF-α) like ligand for the drosophila EGFR gurken (grk), and expression of spitz in miR-8 positive glia is required for neuroepithelial cell expansion, neuroblast generation, and thereby normal optic lobe development.(*154*) Data from our lab, as well as from others, demonstrate that spatio-temporaly regulated expression of the miRNA let-7 controls neuronal differentiation, and is required for normal MB formation. The insect hormone ecdysone regulates many aspects of development, and pulses of

ecdysone that occur during larval and pupal stages initiate major developmental programs. In the developing brain, temporally controlled ecdysone signaling drives let-7 expression, which in turn negatively regulates the transcription factor abrupt (ab). Abrupt negatively regulates the cell adhesion molecule fasciclin II (FasII). FasII is expressed in the α/β and γ lobes of the MB, but not in the α'/β' lobes. Tight regulation of FasII expression is required for normal MB development, and ab mutants exhibit ectopic FasII expression in the α'/β' lobes. Expression of FasII in α'/β' lobes results in α'/β' neurons invading the α/β lobes, and conversely expression of ab in α/β lobes leads to innervation of α'/β' lobes by α/β neurons. α/β neurons are the last MB neurons generated during the pre-pupal to pupal transition, a period corresponding to elevated ecdysone and let-7 levels. Decreasing levels of the transcription factor chronically inappropriate morphogenesis (chinmo), in post-mitotic MB neurons specifies the sequential differentiation of each of the MB cell types. Chinmo is a let-7 target, and the decline in its expression corresponds to increasing let-7 expression. Thus, let-7 controls neuronal differentiation and wiring of the *Drosophila* brain.(*155*, *156*) MicroRNAs also play a vital role in modulating synaptic efficacy. There are now numerous examples in many organisms of microRNA mediated mechanisms that regulate neuronal function. The microRNAs of the *Drosophila* miR-310, miR-311, miR-312, and miR-313 cluster (mir-310-313) negatively regulate kinesin heavy chain 73 (Khc-73) through target sites in its 3'UTR.(*70*) Kinesins are involved in transport of various cargoes to synapses, and Khc-73 is involved in transport along microtubules and interacts with Rab-5 containing vesicles.(*157*, *158*) Loss of mir-310-313 in *Drosophila* motor neurons

results in increased Khc-73 expession, accumulation of the presynaptic active zone marker Bruchpilot (brp), and increased calcium influx. In this way, mir-310-313 negatively regulates the efficacy of neuro-muscular junctions (NMJ) in *Drosophila* motor neurons by reducing Khc-73 mediated transport.(*70*, *157*, *159*) In an example of the direct role microRNAs play in learning and memory, expression of the *Drosophila* miR-276a in the MB is required for LTM formation. Dopaminergic signaling in the MB via the dopamine receptor DopR is also required for LTM formation. miR-276a regulates DopR, and removing a single copy of the DopR gene rescues LTM defects in miR-276a mutants. miR-276a acts by tuning DopR levels, and this regulatory pairing is required for LTM.(*160*) MicroRNA regulation of synaptic efficacy is dynamic, and responds to neuronal signaling. In *Aplysia*, the CNS enriched miR-124 is rapidly down-regulated in response to serotonin induced MAPK signalling, and does not return to baseline levels for 12 hours. miR-124 negatively regulates CREB, and reduced silencing of CREB by miR-124 results in lasting enhancement of synaptic efficacy.(*161*) As CREB is a transcriptional regulator, reduced translational silencing by miR-124 initiates lasting changes in expression of many genes beyond the set it directly silences, even after it returns to baseline levels. More recently, Nesler et al reported that chronic activation of motor neurons in *Drosophila* larvae induces changes in expression levels of several miRNAs. They also report that these changes result in alterations in synaptic size and number at the NMJ.(*162*) However, chronic activation does not resemble normal patterns of activity in the CNS, and NMJs differ from synapses in the brain in many important ways. Therefore, this result may not reflect changes in miRNA

expression induced by activity relevant to memory formation. Nonetheless, this result indicates that neural activity can induce complex changes in microRNA expression patterns, and that examination of single microRNAs may not be sufficient to understand the regulatory contribution of miRNAs in response to neural activity. Neurons have perhaps the most elaborate morphology of any cell type, and can extend processes to great lengths, even a meter or more. Rapid and lasting changes in synaptic plasticity would be unachievable in such cases if it were to require transport to synapses of proteins newly synthesized in the soma. As such, localized synthesis of proteins is a particularly important feature of neuronal physiology. Neurons localize certain mRNAs and microRNAs to distinct cellular regions, including at synapses. Dendrites and axons contain different sets of translationally silenced mRNAs, poised for local protein production in response to appropriate patterns of neural activity. However, while a large complement of mRNAs is present at dendrites, they are largely excluded from axons. Preferential localization of mRNAs within dendrites reflects the importance of postsynaptic local translation. Indeed, translation within dendrites isolated from the soma can effect lasting changes in synaptic efficacy. Such local translation enables modulation of synaptic efficacy within selected dendritic branches or even at individual synapses, while leaving more distant synapses within the same cell unchanged. Synapse specific, or dendritic branch specific plasticity is thought to play a vital role in memory formation, as it allows a neuron to selectively tune the strength of those synapses receiving learning relevant input, while maintaining relatively stable connectivity with its other uninvolved synaptic partners. (*163-167*) In neurons, mRNAs are

selectively transported to specific cellular locations by forming ribonucleoprotein (RNP) complexes with RNA binding proteins (RBPs). These RNPs are incorporated into transport granules, and trafficked to their destinations along microtubules. RBPs involved in localizing mRNAs within neurons largely depend on interactions with sequence elements present in the 3'UTR. Transport granules also contain RISC components, and some RBPs involved in mRNA localization physically interact with RISC component proteins. Interactions between mRNAs and their RBP partners, as well as interactions between RBPs are modulated by neuronal activity.(*153*) Some mRNAs are delivered to dendrites in a translationally silenced state. When protein synthesis is dictated by synaptic activity, local signaling leads to covalent modification of RBPs, and subsequent changes in translation of proximal mRNAs. However, our understanding of how synaptic activity is transduced into local changes protein synthesis remains incomplete. Results from several studies indicate that microRNA mediated silencing controls local translation of specific mRNAs at or near synapses, and that this silencing can be modulated in response to neural activity. A recent study in the rat brain demonstrated that activity induced changes in miRNA expression do not directly translate into changes in levels of miRISC bound miRNAs.(*168*) This suggests that association of microRNAs with argonaute proteins is itself regulated in response to neuronal activity in a way that is independent from changes in miRNA biosynthesis. However, this study did not examine whether microRNA-argonaute association is controlled by changes in miRISC loading or microRNA degradation. Interestingly, certain pre-miRNAs and the microRNA biogenesis factor Dicer, are present in rat hippocampal

synaptoneurosomes. In rats, pre-miR-134 is localized to dendrites through the interaction of sequences present in its terminal loop with the DEAH-box helicase DHX36. Further, silencing of target mRNAs by miR-134 in dendrites is DHX36 dependent, and knockdown of DHX36 leads to enlarged dendritic spines.(*169*) miR-134 silences the LIM domain containing kinase LimK1, and loss of LimK1 mRNA silencing by miR-134 leads to increased dendritic spine size.(*170*) It is therefore possible that at least some pre-microRNAs are processed into mature microRNAs in dendrites, and that regulation of microRNA processing provides a control point for modulating microRNA mediated translational silencing near synapses. Additionally, microRNA and miRISC interactions with targeted mRNAs are modulated by signaling pathways known to control synaptic plasticity. Experiments in our lab, and by Banerjee et al provide evidence that local degradation of miRISC components at synapses in response to neuronal activity leads to synthesis of proteins involved in synaptic plasticity. (*126*, *171*) In mice, the RISC component MOV10 is required for translational silencing of CaMKIIα, Limk1, and Lypla1. MOV10 is degraded at synapses in a proteasome dependent manner in response to NMDA receptor signaling. Using a fluorescent reporter of protein synthesis harboring the Lypl1a 3'UTR, the authors showed that degradation of MOV10 corresponds to new synthesis of the reporter at synaptic sites following NMDA receptor stimulation.(*126*) Subsequent work showed that degradation of proteins in the rat amygdala via the ubiquitin-proteasome (UPS) pathway is required for long term fear memory, and that NMDA receptor signaling during memory formation and retrieval triggers UPS mediated degradation of MOV10.(*172*) A recent study of

MOV10's interaction with mRNAs and other RISC component proteins indicates that MOV10 functions as an RNA helicase that facilitates UPF1 mediated 5' to 3' exonucleolytic decay via XRN1 and XRN2 by unwinding secondary structures that inhibit such decay.(*173*) 5' to 3' exonucleolytic decay via XRN1 is a known feature of miRNA mediated silencing. MicroRNA mediated silencing may also be controlled in ways other than regulated destruction of miRISC components. MicroRNA-target binding is regulated by RBPs aside from miRISC components. In mice, NMDA receptor signaling leads to activation of the kinase mTORC1, which in turn inhibits binding of the RBP HuD to mRNA encoding the voltage gated potassium channel Kv1.1. In the absence of HuD, miR-129 silences Kv1.1 mRNA. Without mTORC1 activity, HuD binds Kv1.1 mRNA, and thereby relieves it from miR-129 induced silencing.(*174*) Similarly, the HuD related protein HuR relieves target mRNAs from silencing by let-7. HuR binds sequences in the 3'UTR, and this binding results in displacement of miRISC from the target mRNA. Sequences bound by HuR can be tens of bases distant from microRNA target sites, and relief from microRNA mediated silencing by HuR is dependent on its ability to oligomerize. Further, HuR is unable to displace miRISC proteins that are directly tethered to mRNAs, and the ability of HuR to displace miRISC is reduced if microRNA-target complementarity is perfect. These results indicate that HuR, and perhaps other similar proteins such as HuD and the *Drosophila* protein ELAV, relieve microRNA mediated silencing by interfering with basepairing of microRNAs with their targets.(*175*) The activity of ELAV family proteins including HuD and HuR has been shown to be modulated by covalent modification and proteolytic cleavage. (*174*, *176*, *177*) Thus, ELAV family

proteins may be a major factor in transducing synaptic signaling into changes in microRNA mediated translational silencing. Several RBPs have been implicated in control of microRNA mediated translational control at synaptic sites. For instance, fragile X syndrome, the most common monogenetic form of intellectual disability is caused by loss of the RBP FMRP. FMRP is found at synapses, and is involved in localizing mRNAs to dendrites and in control of dendritic protein synthesis. In its phosphorylated state, FMRP associates with the mammalian microRNA effector argonaute protein, and leads to formation of a FMRP-miR-125a-miRISC that silences mRNA encoding the postsynaptic density protein 95 (PSD-95). Signaling from metabotropic glutamate receptors (mGluR) triggers dephosphorylation of FMRP, subsequent dissociation of FMRP-miR-125a-miRISC from PSD-95 mRNA, and leads to rapid translation of PSD-95.(*178*) Given the many RNA-protein and protein-protein interactions that occur along the 3'UTR of mRNAs, it is likely that additional mechanisms controlling miRNA mediated silencing in response to neuronal activity will continue to be discovered for some time to come. However, existing work has demonstrated that reversible silencing of mRNAs by microRNAs is an important feature of translational control at or near synapses, and is involved in lasting forms of memory. Intriguingly, recent work demonstrates that microRNAs also act via multiple mechanisms as intercellular signaling molecules. MicroRNAs are packaged into extracellular vesicles, and are secreted by cells. Several microRNAs, including let-7b, are ligands for the toll like receptor TLR7. TLR7 can directly activate the transient receptor potential cation channel TRPA1. TLR7 stimulation with let-7b induces inward currents and action potentials in cells coexpressing TLR7 and

TRPA1. TRPA1 is expressed in nociceptive neurons in mice, and its activation causes pain. Injected and endogenous intercellular let-7b triggers a pain response that is sequence dependent, and can be blocked by inhibitors of TRPA1 or pretreatment with let-7b inhibitors. Interestingly, let-7b is highly expressed in nociceptive neurons in mice, and is secreted in response to neuronal activity in these cells.(*179*) Thus, let-7b and other microRNAs harboring a specific shared nucleotide sequence act as signaling molecules in neurons. Secreted microRNAs can also be taken up by recipient cells where they are capable of silencing mRNAs.(*180*) MicroRNAs released in exosomes at synaptic sites in response to neuronal activity are taken up by glia and can alter expression of genes involved in regulating neurotransmitter levels in the synaptic cleft, thereby modulating synaptic efficacy.(*181*) While microRNAs are known to be essential for many biological processes, and in virtually all cell types in eukaryotes, they have uniquely diverse and vital functions in neurons. Underscoring the importance of microRNAs in proper function of the nervous system are numerous examples of intellectual disability and neuromuscular diseases caused by mutations in genes that encode miRISC interacting proteins. Moreover, the pathologies and symptoms of diseases caused by mutations in these genes are strongly conserved from insects to humans, demonstrating the ancient and central function of the microRNA pathway in the nervous system. However, in order to fully understand the contributions to neuronal function of microRNAs, additional studies examining the activity of all microRNAs simultaneously are needed. Previous work has studied the plasticity relevant dynamics of individual microRNAs in several model animal systems, but

the full spectrum of microRNA regulation during learning and memory formation remains largely unstudied and is poorly understood. Thus, the importance of microRNAs in synaptic plasticity and LTM is well established, but the complex interplay of posttranscriptional regulation by hundreds of microRNAs requires additional study to fully understand their function in these processes.

esiRNAs, piRNAs, And Novel sRNAs In Neurophisiology

Studies of sRNA activity in the nervous system have thus far focused almost exclusively on microRNAs, and have left the potential roles of other classes of sRNAs in neurons largely unexplored. The initial discovery of piRNAs in reproductive tissues at first biased the exploration of their function away from somatic cells, let alone the nervous system. More recently, piRNAs have been shown to be well expressed in a variety of somatic tissues from several species.(*58*, *60*, *61*, *137*) Neurons of the Drosophila brain express piRNAs and the piRISC components Aub and Ago3. Interestingly, expression of Aub and Ago3 is lower in αβ neurons of the MB than in neighboring MB neurons. The most well studied function of the piRNA pathway is in maintaining genome integrity by posttranscriptionally silencing transposon expression. Accordingly, transposon expression is elevated in αβ neurons. Expression of retrotransposons can lead to their mobilization, and formation of genomic lesions at the sites of excision and disruption of genes at insertion sites. In αβ neurons, de novo transposon insertion into exons preferentially occurs within genes annotated with neural gene ontology (GO) terms.

Insertion into promoter regions occurs largely within promoters that are active in αβ neurons. These findings indicate that preferential reduction of piRNA pathway activity within neurons required for LTM drives genetic diversity amongst these cells, and likely results in differences in their physiology.(*59*) piRNAs may also play an important role in activity dependent changes in gene regulation. piRNAs are found in the Aplysia CNS, including piRNAs matching a CpG island in the promoter of the CREB2 gene. Following exposure of CNS neurons to serotonin, the CREB2 promoter matching piRNAs are induced. Consistent with their known ability to transcriptionally silence genomic regions with which they are complementary, induction of these piRNAs leads to methylation of the CREB2 promoter, and a resulting piwi dependent reduction in CREB2 expression. CREB2 inhibits CREB1 driven transcription, thereby negatively regulating memory formation.(*60*) This finding therefore implicates the piRNA pathway in transcriptional regulation involved in lasting memory formation. No other example of such a piRNA mediated mechanism controlling gene expression in response to memory relevant signaling has been published to date. It is possible that the mechanism may be unique to transcription dependent memory in *Aplysia*, but the dearth of published studies examining piRNA expression in the brain leaves open the possibility of continuing discoveries in this vein. While reports of piRNA involvement in neurophysiology are few and far between, studies of esiRNA involvement are rarer still. However, esiRNAs are expressed in the CNS of nematodes, insects, and mammals. Though research into the functions of esiRNAs in neurons is in its infancy, evidence supporting their involvement in processes relevant to memory formation already

exists.(*72*, *113*, *114*, *182*) In *C. elegans*, the guanylyl cyclase ODR-1 is required for odor sensation. Adaptation, a process in which prolonged exposure to a given odorant while undergoing nutritional starvation results in decreased atraction to that odorant, requires downregulation of ODR-1.(*183*) Odor adaptation also requires expression of the siRNA effector argonaute protein NRDE-3 in olfactory sensory neurons. esiRNAs complementary to odr-1 coimmunoprecipitate with NRDE-3 in these cells. esiRNAs matching odr-1 are upregulated following odor adaptation, and this upregulation corresponds with a decrease in odr-1 mRNA. esiRNAs are thought to act cotranscriptionally, and exert silencing through heterochromatin formation. Following odor adaptation, the heterochromatin binding protein HPL-2 is found at the odr-1 gene locus. Thus, it appears that prolonged stimulation of olfactory sensory neurons leads to production of esiRNAs targeting odr-1, a gene required for olfactory sensation. This in turn results in heterochromatin formation at the odr-1 locus, reduced ODR-1 protein expression, and lasting attenuated sensitivity to odor.(*72*) The mouse hippocampus produces a variety of sRNAs, including esiRNAs that map to predicted hairpin RNA forming regions within genes. Interestingly, genes annotated with the GO term synapse are overrepresented in this set. Following olfactory discrimination training of mice, hippocampal expression esiRNAs as a class, and sRNAs derived from snoRNAs and other non-coding RNAs increased.(*71*) However, this study did not use statistical analyses sophisticated enough to determine whether the expression esiRNAs from any given locus changed following training. Nor did it examine expression of any of the putative mRNA targets of the esiRNAs identified. As few studies examining

esiRNA function in synaptic plasticity or memory have been published, the existing evidence does not conclusively support or reject the existence of such a mechanism. The demonstration of esiRNA involvement in nematode odor adaptation is encouraging, and argues for further study of the possibility that esiRNAs are involved in synaptic plasticity or memory in higher animals, though significant differences between nematode, mammalian, and insect esiRNA pathways exist. A similar situation exists in the case of piRNAs, with the tantalizing possibility that CREB2 regulation by piRNAs in mollusks is only one example of a broader mechanism for formation of stabile memory. Furthermore, the recent advent and continuing improvement of high throughput sequencing methods has revealed a huge array of other previously unknown sRNAs. Little is known about the biological significance of most of these novel sRNAs, and it is possible that some may be involved in memory formation. Unbiased surveys of sRNA expression during memory formation remain lacking, and could provide significant insight into the ways in which gene expression is controlled during this process.

Memory formation is the result of alterations in patterns of activity in networks of neurons. Such alterations are achieved through making, breaking, and tuning the efficacy of synaptic connections between neurons. While fleeting memories can result from covalent modification of existing proteins at synapses, lasting memories are formed through regulated production of new proteins, and changes in gene expression programs in neurons. MicroRNAs have emerged as central players in neuronal function, governing expression of genes in key pathways

involved in synaptic plasticity and memory formation. In addition to their essential activities in the soma, mounting evidence indicates that reversible silencing of translation by microRNAs in dendrites is an important feature of stable memory formation. To date, most research into the function of microRNAs in neurons has focused on individual microRNA-target regulatory pairings. However, as an individual microRNA may have many mRNA targets, and an individual mRNA may be targeted by multiple microRNAs, it is clear that a more systems based approach is needed. The recent advent and rapid evolution of high throughput sequencing technologies has greatly expanded our understanding of microRNA genetics, biogenesis, function, and regulation. Using this technology, one can measure the expression of all microRNAs simultaneously, even in very small tissue samples. Further, unlike other technologies such as RT-PCR or hybridization, high throughput sequencing does not require a priori selection of the sequences to be studied. This feature of high throughput sequencing has lead to the rapid expansion of the catalog of known microRNAs, and uncovered unexpected aspects of their biology. It has also unveiled a great diversity of previously unknown functional sRNAs including piRNAs and esiRNAs. The regulatory potential of these new classes of sRNAs is only beginning to be understood, and substantial opportunities for discoveries in this area exist, particularly in the context of synaptic plasticity and memory formation. In chapter II, I will describe experiments directed at understanding how sRNAs are regulated during formation of long-term memory in *Drosophila,* and identify several microRNAs, piRNAs, and esiRNAs exhibiting a response to aversive olfactory conditioning.

Part III: The Beta Secretase Beta-Site APP-Cleaving Enzyme (BACE) In Memory Formation and Cognitive Impairment

Proteolytic Processing Of Amyloid Precursor Protein (APP) Family Proteins Is Involved In Alzheimer's Disease Pathology

Our experience of life occurs largely in retrospect, with virtually all that happens to us immediately contextualized with what has come before. Whether this takes the form of a whiff of fragrance that quickly brings to mind a narrative recollection of an important time with a loved one, or in the form of our ability to unthinkingly find our way home from the office, the past is constantly brought into the present and often dictates our actions. Remembering skills acquired in the past and faithfully repeating the actions they involve is perhaps the greatest evolutionary advantage humans possess. Our ability to form supportive social groups is our other major evolutionary advantage, and our navigation of these groups largely defines modern life. Social interactions are entirely dependent on being able to learn new names and associate them with faces in the future, to recall past conversations, details of shared experience, and to remember facts about the people one meets. Accordingly, injuries or diseases that disrupt memory acquisition and retrieval are amongst the most crippling afflictions for those affected and are a great burden for their caretakers. Alzheimer's disease (AD) and related neurodegenerative dementias are among the most common late onset cognitive

impairments, affecting tens of millions of people world wide, and are predicted to increase in prevalence as life expectancies rise in developing countries. Substantial effort has been directed at understanding the disease mechanisms of these conditions. Much of what we know about AD comes from studies of brain tissue of patients from families exhibiting heritable susceptibility to early onset of the disease. This work has described features associated with disease progression. In AD, the microtubule associated protein tau (Tau), which is involved in regulating axonal transport, becomes hyperphosphorylated in neurons, and this results in formation of Tau aggregates termed tangles. Accumulation of these tangles disrupts transport along microtubules, and results in cell death. This condition is termed tauopathy, and is a shared feature of neurodegenerative diseases other than AD. As AD has unique aspects, and dementia is not a feature of all diseases featuring tauopathy, it is logical to assume that tauopathy is a result, and not the root cause of AD. Progression of AD is also associated with formation of extracellular senile plaques near synaptic sites. Senile plaques are largely composed of aggregates of a single 42-43 amino acid peptide termed amyloid-β (Aβ), which is derived from cleavage of the amyloid precursor protein (APP) by one of two beta-site APP cleaving enzymes (BACE1 and BACE2). As AD is strongly linked to trisomy of chromosome 21 in humans, APP was first identified by screening cDNA libraries generated from brain tissue collected from individuals with this genetic abnormality using oligonucleotide probes generated based upon the sequence of Aβ.(*184*) This approach lead to the identification of an mRNA encoding a 695 amino acid transmembrane protein. Subsequent work showed that Aβ is derived from a region

near the C-terminus that includes extracellular and intramembrane portions of the APP sequence. APP is sequentially cleaved by multiple proteases, yielding several functional peptides. APP processing occurs in one of two pathways that are distinguished by the initial cleavage event, and yield metabolites with distinct activities. The Aβ peptide is the product of processing initiated by cleavage of APP by BACE1 or BACE2, membrane-bound aspartic proteinases termed β-secretases. This cleavage defines the N-terminus of Aβ, and liberates the soluble extracellular N-terminal domain of APP termed $sAPP^{\beta}$. The 99aa membrane bound C-terminal fragment produced by β-secretase cleavage is termed C99 and includes the Aβ sequence. Subsequently, C99 is cleaved within the membrane by a membrane-bound aspartic proteinase termed γ-secretase. This second cleavage event yields Aβ, and a C-terminal intracellular domain fragment (AICD). Aβ can move into extracellular space, where it can lead to senile plaque formation. AICD is transported to the soma and enters the nucleus, where it can act as a transcriptional regulator. However, β-secretase initiated APP processing is the exception, and most APP processing is instead initiated by ADAM10, an α-secretase type zinc metalloproteinase. Cleavage by the α-secretase occurs within the Aβ region, and liberates the extracellular N-terminal domain and part of the Aβ sequence from the cell surface. This metabolite is termed soluble APPα (sAPPα). As cleavage by α-secretase occurs within the Aβ sequence, Aβ generation is precluded. The 83aa α-secretase generated C-terminal fragment is termed C83, and like C99, is cleaved within the membrane by γ-secretase. This second cleavage event yields a short membrane bound peptide termed P3, and AICD. However, unlike AICD production

initiated by β-secretase cleavage of APP, AICD produced as a result of cleavage by α-secretase is largely excluded from the nucleus and degraded, and does not induce the same transcriptional changes as β-secretase derived AICD.(*185*) AICD is rapidly degraded in the cytosol, but BACE1 is largely found in recycling endosomes, and trafficking to the nucleus via the endosomal pathway protects AICD, thereby allowing it to regulate transcription. (*186-188*) The coupling of Aβ production and AICD mediated changes in transcription suggests the possibility that accumulation of senile plaques may not in itself drive AD, and is instead a symptom of misregulated β-secretase activity and downstream signalling. This view is supported by the observation that cognitive decline in human AD is not halted even when senile plaques are cleared following administration of monoclonal antibodies directed against Aβ. Further, Aβ plaques can be found in brains from non-demented patients, and dementia can occur in the absence of significant Aβ plaque accumulation.(*189*) Work in mammalian model systems further complicates the picture. For instance, endogenous APP and Aβ is required for LTP in the mouse hippocampus, indicating that Aβ production is a function of normal neurophisiology.(*190*) Indeed, though there is substantial evidence implicating the amyloidogenic pathway as causal in AD, and despite significant efforts in academic and industry labs, no therapy targeting this pathway has yet proven successful enough for use in treating AD in humans. This may be in part due to the extensive use of cell culture and in vitro methods to study the function of APP and its metabolites. Numerous differences between in vitro and in vivo results in studies of APP processing and function have been documented. This fact is indicative of the

need for studies to be conducted in animals in order to better understand the role of APP and its processing in normal neurophysiology, and in AD. As APP cleavage by BACE generates Aβ, and has been shown to be required for LTP, mechanisms regulating APP cleavage by BACE, and events downstream of this cleavage that regulate synaptic plasticity are of particular interest.

APPL Is The Drosophila Homologue Of APP, And Its Processing And Functions Are Conserved

Recently, *Drosophila* has been developed as a model organism in which to study AD. The relative simplicity and accessibility of the *Drosophila* brain, as well as the ease of measuring memory formation and retention in large numbers flies using behavioral studies, makes experimental approaches possible that would be prohibitively time consuming or difficult in mammals. Further, the availability of existing reagents with which to genetically dissect AD in flies makes rapid advances in our understanding of the underlying disease mechanisms possible. APPL, The *Drosophila* homologue of APP, is processed similarly to APP, and when human APP is expressed in flies, it too is processed much as it would be in humans. As is the case in mammals, processing of APP and APPL in *Drosophila* can occur along either an α-secretase or β-secretase initiated pathway. The α secretase pathway involves Kuzbanian (Kuz) and Presenillin (Psn), homologues of α and γ-secretases respectively. (*191-195*) Kuz and Psn have been extensively studied in the context of proteolytic processing of the *Drosophila* signaling protein Notch, which also involves

sequential cleavage by α and γ-secretases. Similar to AICD production from APP, Notch cleavage also produces a transcription regulating C-terminal fragment, dubbed NICD. The Notch pathway is one of the most well studied systems regulating growth and development, and thus provides an excellent intellectual scaffold upon which to build, and a large set of tools with which to work.(*195*) More recently, a homologue of the human BACE genes (dBACE) was identified in *Drosophila*. While mammals possess two BACE genes (BACE1 and BACE2), dBACE is the only *Drosophila* homologue identified thus far. Cleavage of APPL by Kuz or dBACE occurs within the extracellular domain, and liberates a large soluble fragment (sAPPL). Both Kuz and dBACE cleave at positions toward the C-terminal end of APPL, near the cell surface. While the human α- cleavage site of APP is closer to the C-terminus than the β-cleavage site, the situation is reversed in Drosophila, with dBACE cleaving closer to the C-terminus than Kuz.(*194*) Human APP expressed in *Drosophila* cells is cleaved by dBACE, and this cleavage initiates amyloidogenic processing similar to that occurring in humans. Mirroring AD pathology, transgenic expression of human APP or Aβ in the fly brain leads to plaque formation, neurodegeneration, and behavioral defects. (*193*, *194*, *196*) Though the Aβ sequence is not conserved in APPL, aged flies overexpressing APPL and young flies co-overexpressing APPL and dBACE in the brain exhibit formation of plaques that are Thioflavin-S positive, a classic marker of senile plaques in mammals. Plaque formation in flies induced by pan-neuronal co-overexpression of APPL and dBACE is associated with behavioral defects that worsen with age.(*194*) The relatively late discovery of dBACE long discouraged use of Drosophila as a model organism in

which to study the role of APP processing in synaptic plasticity, neurodegeneration, memory, and behavioral defects. However, with the addition of dBACE to the list of conserved elements of the APPL processing machinery, it is now clear that the complete pathway is evolutionarily ancient, and that work in flies can provide unique contributions to our understanding memory formation and AD.(reviewed in (*197*) and (*198*))

Proteolytic APPL Processing

APPL and its homologues are expressed in many tissues, and are involved in regulating diverse processes including growth, calcium, insulin and glucose homeostasis, apoptosis, mitochondrial function, and synaptic plasticity. (*199-201*) This conserved family of transmembrane proteins appear to act as both receptors and ligands in these processes, interacting with other cells through their extracellular domains, and transducing signals to the nucleus through their intracellular domain. Proteolytic processing of APPL is a major aspect of its activity, and generates several functional metabolites.(reviewed in (*197*, *200*)) The extracellular domain of APPL includes 2 regions of extensive conservation dubbed E1 and E2. The E1 region is closest to the N-terminus, and includes a heparin-binding/growth-factor-like domain, and copper and zinc binding domains, and may be involved in functional interactions such as dimerization, ligand binding, and transfer of metal ions to other proteins. The E2 conserved region dimerizes in solution, such that the N-terminal end of one E2 region packs with the C-terminal

end of the other. Thus, the conserved structures within the extracellular domain of APPL suggest that it is functional, and is involved in protein-protein interactions.(*202*, *203*) As previously mentioned, the α and β cleavage sites lie just outside of the cell surface, and cleavage by dBACE or Kuz liberate nearly the entire extracellular domain (sAPPL$^{\beta}$ and sAPPL$^{\alpha}$, respectively). The resulting C-terminal fragments are ~100aa long, and include the transmembrane domain, and a highly conserved intracellular domain. The transmembrane domain contains the γ cleavage site, which is cleaved by Psn. γ cleavage by Psn yields a highly conserved AICD and an unconserved short peptide whose N-terminus is defined by either the α or β cleavage site. The β-cleavage derived peptide (dAβ), though dissimilar to Aβ in sequence, forms toxic aggregates, suggesting that it is functionally similar to Aβ. The AICD also includes a $G_o\alpha$-binding domain, an internalization signal, and a conserved YENPTY motif, which mediates binding with the adapter proteins X11/Mint and Disabled (Dab).

Though the presence of AICD has been demonstrated in *Drosophila*, its functions remain unkown.(*195*) However, the extensive sequence conservation within the intracellular region of APPL and the similarity of APPL processing and function to that of APP suggest strongly that *Drosophila* AICD functions analogously to human AICD. (*194*, *200*, *204*, *205*) In humans, APP and BACE-1 are trafficked to synaptic sites in separate vesicle types, and remain segregated until appropriate signaling triggers their convergence. BACE-1 is localized largely to recycling endosomes, and thus prevented from interacting with APP. AICD contains an internalization signal that is required for endocytosis of APP, and for colocalization of APP and BACE-1.

Mutation of the APP internalization signal, or blockade of endocytosis also reduce Aβ production.(*206*) BACE-1 activity is optimal in an acidic environment such as that found within endosomes. Furthermore, the γ-secretase cleavage site differs when cleavage occurs at the plasma membrane vs. in endosomes. Thus, BACE-1 initiated processing of APP can be regulated independently of ADAM10 initiated processing through endocytosis of APP.(*206*, *207*) Such a mechanism could explain why nuclear localization of AICD is largely driven by β-secretase initiated processing. Though segregation into recycling endosomes of dBACE initiated processing of APPL has not yet been demonstrated, the functional conservation of dBACE suggests strongly that this mechanism is also conserved.

APPL And Its Metabolites Are Involved In Neurodevelopment, And Regulate Synaptic Structure

Proper APPL expression is required for normal synapses in *Drosophila*. Synapses can form in flies carrying a mutation that eliminates APPL expression ($APPL^{d}$), but these flies produce fewer synaptic boutons at larval NMJs, while flies overexpressing APPL exhibit increased synaptic bouton number and defects in synaptic bouton size at larval NMJs. Expression in *Drosophila* neurons of APPL processing products or of APPL transgenes carrying mutations that affect the production and function of these processing products, shows that APPL processing regulates synaptic number, structure, and function through several pathways.(*204*) The initial step in APPL processing is cleavage by either Kuz or dBACE, and liberates

the corresponding N-terminal fragments sAPPL$^{\alpha}$ or sAPPL$^{\beta}$ from the membrane. Conserved domains within the extracellular region of APPL suggest that, once liberated from the cell surface, it may function as a diffusible ligand. Alternatively, cleavage of APPL by Kuz or dBACE could be the functional event, initiating intracellular signaling events, and extracellular APPL fragment production only an intermediate step on the way to its degradation following cleavage. Torroja et al. found that expression of a transgene encoding the APPL extracellular domain, and with a C-terminus that does not include the α or β cleavage sites (sAPPL), in larval motor neurons is hampered by rapid turnover, and produces no phenotype at the NMJ unless expression is strongly driven by multiple copies of the transgene. Even when expression is driven at high levels, no change in the number of synaptic boutons is evident, though the number of satellite boutons is reduced. Conversely, expression of a transgene encoding APPL with both α and β cleavage sites deleted (APPL^{sd}) mimics the phenotype produced by overexpression of APPL, and yields increased numbers of both satellite and parent boutons. Further deletion of either the E1 or E2 regions of APPL^{sd} (APPL$^{sd\Delta E1}$ and APPL$^{sd\Delta E2}$ respectively) eliminates the increase in parent boutons produced by APPL^{sd} expression, but still yields increased satellite boutons. However, expression of APPL^{sd} mutated such that the C-terminal intracellular domain is deleted (APPL$^{sd\Delta C}$) produces no change from wild type in synaptic structure at the NMJ.(*192*, *204*, *208*) These results indicate that APPL influences NMJ structure largely through its highly conserved AICD, though the extracellular domain contains activities that can influence AICD mediated control of NMJ structure. APPL also regulates neurite outgrowth in cultured *Drosophila*

neurons. Curiosly, cultured neurons from flies overexpressing APPL, and from APPLd flies, both exhibit reduced neurite outgrowth.(*209*) Functional dissection of the APPL protein in cultured neurons from flies expressing APPL variant transgenes elucidates this result. Cultured neurons from $APPL^{sd}$ or $APPL^{sd\Delta C}$ flies exhibit excess neurite branching and protuberances resulting from actin filament and microtubule abnormalities. Neurites of sAPPL neurons branch less than do those of wild type neurons. Coculture of neurons expressing sAPPL and $APPL^{sd}$ decreases neurite branching in $APPL^{sd}$ cells and reduces the number of $APPL^{sd}$ cells with protuberances.(*209*) This suggest that membrane bound APPL can act as a receptor that positively regulates neurite outgrowth and branching, and that sAPPL is a ligand that inhibits this activity. If so, overexpression of APPL leads to increased sAPPL levels, and this inhibits any signaling that increases neurite outgrown, even when extra APPL is present.(*197*) APPL has also been shown to regulate neuronal morphology in the adult brain. Overexpression of APPL in neurons of the adult brain results in dramatically increased axonal arborization. This increase requires the presence of the conserved YENPTY motif within AICD, the adapter protein Dab, and activation of the Dab interacting dAbl kinase. dAbl is known to regulate the cytoskeleton, and is involved in axon guidance.(*210*) APPL and the products of its proteolytic processing influence numerous aspects of neuronal structure and function. The regulatory functions of APPL processing products also appear to interact in complex ways, and sometimes counteract each other. APPL and its metabolites are able to modulate several signaling pathways, and can alter transcriptional programs. It is therefore possible for changes in the regulation of

the proteolytic events involved in APPL processing to have profound consequences for neuronal connectivity, function, and ultimately behavior.

APPL Processing Regulates Neuronal Activity.

APPL regulates the morphology of neurons, and the number of boutons present at a synapse.(*204*) As synaptic boutons are the sites of neurotransmitter release, their number is related to the efficacy of presynaptic neurons. However, soluble APPL metabolites can also modulate the excitability of neurons by altering ion channels. The presence of sAPPL in cell culture media reduces excitability of cultured *Drosophila* neurons by enhancing K^+ currents. Conversely, spontaneous excitatory postsynaptic potentials (EPSP) are enhanced in neurons expressing $APPL^{sd}$ or $APPL^{sd\Delta C}$. Coculture of neurons expressing sAPPL and $APPL^{sd}$ restores excitability to near wild type levels in $APPL^{sd}$ cells.(*209*) In mammals, $sAPP^{\alpha}$ activates K^+ channels and suppresses NMDAR currents via PKG, thereby reducing excitability. (*211*, *212*) These findings indicate that sAPPL can decrease neuronal activity by enhancing K^+ channel activity, and that sAPPL production can provide negative feedback against increased excitability that is driven by membrane bound APPL. Though the conservation of many aspects of APPL function suggest that $sAPPL^{\alpha}$ is responsible for enhancing K^+ currents in flies, reliance on the sAPPL transgene in work thus far does not allow the relative importance of $sAPPL^{\alpha}$ or $sAPPL^{\beta}$ in modulating neuronal activity to be evaluated. However, in mammals, the

C-terminal end of $sAPP^{\alpha}$, which differs from that of $sAPP^{\beta}$, is required for these activities.(*211*)

Aβ also regulates excitability of neurons in mammals, though the underlying mechanisms differ from those involving $sAPP^{\alpha}$. Both acute and chronic application of physiological concentrations of Aβ reduce the excitability of neurons of the rat prefrontal cortex. Interestingly, while application of high concentration Aβ initially reduces excitability, prolonged exposure enhances excitability, and can lead to tonic firing.(*213*) As is the case with $sAPP^{\alpha}$, Aβ's effects may result from altered K^{+} channel activity and expression. However, Aβ also directly binds and alters the behavior of nicotinic acetylcholine receptors (nAchR). (*214*, *215*) Cholinergic signaling is also controlled presynapticaly by Aβ via alterations in K^{+} conductance, leading to reduced acetylcholine release.(*216*) Further, Aβ alters calcium influx via MAPK phosphorylation of L-type voltage dependent Ca^{++} channels.(*217*) Evidence also indicates that Aβ inhibits NMDAR signaling, though this likely occurs downstream of the receptor, and may be related to the presence of APP in the multiprotein NMDAR complex.(*218*, *219*) As dBACE and dAβ were discovered only recently, the role of dAβ in regulation of neuronal activity remains poorly understood. However, given the similarity of phenotypes generated by Aβ and dAβ manipulations, one might expect that dAβ has many of the same effects on neuronal activity as Aβ.

In addition to the rapid, and perhaps direct effects on neuronal activity just discussed, APPL and its homologues regulate synaptic plasticity via slower and more persistent mechanisms. Internalization of APPL family proteins can induce

NMDAR subunit substitution. The C-terminal intracellular domain of APPL family proteins contains sites for interaction with G_o, and numerous adapter proteins. Ligand binding to the extracellular domain of APP alters G_o signaling and other pathways including MAPK, and the Jun N-terminal kinase (JNK) pathway. APP cleavage is not required to activate G_o signaling, but cleaved AICD can do so in the absence of the extracellular domain. Furthmore, transcriptional changes driven by nuclear AICD can alter expression of genes involved in synaptic plasticity, including APP itself, BACE1, and the Aβ degrading enzyme Neprilysin. Transcriptional regulation by AICD can involve the activity of the histone acetyl transferase Tip60. However, transcriptional changes induced by APPL family proteins and their metabolites remain poorly understood.(Reviewed in (*197*, *201*, *207*)) As these findings demonstrate, APPL family proteins and their metabolites are potent regulators of neuronal activity and synaptic plasticity, and act through modulation of a complex set of mechanisms.

Processing Of APPL And Its Homologues Is Regulated By Neuronal Activity

Expression and proteolytic processing of APPL and its homologues can be stimulated in response to a number of conditions including hypoxia, stress, traumatic brain injury, and importantly, neuronal activity.(reviewed in (*197*, *200*, *201*)) As previously discussed, APPL and its metabolites can alter neuronal activity. Thus, changes in APPL processing induced by neuronal activity can in turn alter activity, thereby setting up complex feedforward and feedback loops that can

stabilize or profoundly alter the behavior of neural circuits. Torroja et al showed that intense neuronal activity driven in the hyperexcitible *eag sh* double mutant results in altered APPL localization and perhaps processing.(*204*) Many studies indicate that Aβ is produced in response to various forms of neuronal activity, including electrical stimulation and pharmacological manipulation of cultured neurons, brain slices, and intact brains. APP is processed in response to muscarinic M1 acetylcholine receptor signaling, releasing sAPPα and downregulating Aβ production.(*220-223*) A similar result is observed following exposure of cultured neurons to AMPAR antagonists.(*187*) However, in virtually all of these studies either pathogenically high levels of activity were induced, APP was overexpressed, animals carrying APP mutations driving aberrant APP processing were used, or some combination of these conditions were present.(*224-226*) Such manipulations may not reflect APP processing in any condition relevant to normal brain function or AD. However, studies of APP processing following more careful manipulation of neuronal activity show that the picture is more complex. Prolonged, and perhaps pathogenic NMDAR activity stimulates APP expression and Aβ production in cultured mouse neurons.(*226*) In a seemingly contradictory result, NMDAR antagonists or blockade of $^{Ca2+}$ channels reduce baseline α-secretase driven APP processing. Also, in contrast to prolonged NMDAR stimulation, brief NMDAR stimulation in cultured neurons induces sAPPα release, and reduces Aβ levels in a calcium dependent manner.(*188*) Further complicating the picture, experiments in intact brains differ sharply from those in cell culture. Verges et al. used microdialysis to manipulate NMDAR activity and to measure APP metabolite levels

in interstitial fluid (ISF) of intact brains of awake and freely moving mice. These experiments show that low doses of NMDA trigger Aβ production, while high doses result in decreased Aβ levels in ISF. Both NMDA concentrations trigger synaptic activity, but activate distinct signaling cascades, and drive APP processing down opposing pathways. Preadministration of tetrodotoxin (TTX), which inhibits voltage gated sodium channels, thereby blocking action potentials, reduces basal Aβ levels and prevents Aβ production in response to low dose NMDA. However, preadministration of TTX does not prevent reduction in Aβ resulting from high dose NMDA. This reveals that NMDA induced Aβ production is driven by the events triggered by action potentials, while inhibition of Aβ production driven by high dose NMDA does not. High levels of Ca^{2+} influx via intense NMDAR stimulation activates several second messenger mediated signaling pathways independent of action potentials. Preadministration of compounds that inhibit one such mechanism, the extracellular regulated kinase (ERK) pathway, transforms the response to NMDA administration such that both low and high dose NMDA drive Aβ production. ERK inhibits γ-secretase, and activates α-secretase, thereby reducing Aβ production.(*227*) Thus, while NMDAR signaling induces APP processing, the duration or intensity of NMDAR activity, and cellular setting governs the type of APP processing induced. The view that the balance of α and β cleavage initiated processing is governed by a competition between mechanisms involved in action potentials that drive Aβ production, and ERK driven α-processing, is supported by recent work showing that activity induction leads to clathrin dependent endocytosis of APP and processing by BACE1 in recycling endosomes.(*206*) BACE1 is normally

localized to recycling endosomes, and thus segregated and prevented from acting on APP. Action potentials drive neurotransmitter release via vesicle fusion with the plasma membrane, and clathrin dependent recycling endocytosis. This results in APP internalization, and trafficking such that it is colocalized with BACE1 in pH conditions optimal for BACE1 activity. This leads to β and γ cleavage, and results in AICD production.(*206*) Under certain conditions, neuronal activity can drive intracellular calcium levels high enough to initiate signaling cascades not triggered by neuronal activity at lower calcium concentrations. Such calcium dependent signaling can drive α-cleavage and inhibit γ-cleavage of APP.(*227*) Such processing precludes APP processing that results in Aβ and AICD production. Thus, APP processing is regulated by the level and nature of neuronal signaling. As the various metabolites of APP family proteins exert different effects on neuronal activity, the nature of activity induced APP processing provides important regulatory feedback. Our understanding of the mechanisms governing the regulation of APP processing is incomplete. This is due in large part to the difficulty of work in intact mammalian brains. The biochemical and functional similarity of APPL processing to that of APP may permit discoveries in Drosophila to inform our understanding of these processes in humans. The ease with which neuronal activity and gene expression can be manipulated in selected circuits of intact fly brains could provide a means by which to address remaining questions regarding the regulation of APP processing and its consequences in mammals.

Tight Control Of APPL Expression And Processing Is Required For Drosophila LTM

In addition to the well established role it plays in AD, many studies have demonstrated that APP and its appropriate processing are required for normal memory in mammals.(reviewed in (*201*, *228*, *229*)) Similarly, genetic, biochemical, and behavioral studies in *Drosophila* demonstrate that APPL has major functions in memory as well. APPL is highly expressed in the MB of adult flies, and is required during development for normal neurite morphology in the MB. Manipulations of APPL processing, or expression of human APP fragments in the drosophila brain results in structural defects in the MB.(*209*) Thus, both proper APPL expression and processing are required for normal MB structure. Defects in MB structure or function disrupt olfactory memory formation in *Drosophila*. Therefore, high levels of APPL expression confined to this structure strongly implicate this gene in olfactory memory. Experiments in which siRNAs directed against APPL are inducibly expressed only within the α/β and γ neurons of the adult MB at various time points during several aversive olfactory classical conditioning paradigms show that APPL is not required for learning, STM, or ARM, but is required for LTM. Conditional overexpression of human APP in the adult MB disrupts LTM as well(*230*) Thus, LTM requires that APPL expression in the MB falls within a suitable range. As transcription is also specifically required for LTM, this suggests the intriguing possibility that APPL cleavage leading to AICD mediated transcriptional changes may be a key event in LTM formation. If this is the case, one would expect that Psn and dBACE mutants also affect memory formation. Proper Psn activity is required for courtship memory in male flies.(*231*) Psn acts on targets other than

APPL, including Notch, so it is possible that the memory defects observed in Psn mutant flies are not the result of disrupted APPL processing alone. However, pan-neuronal overexpression of human APP or human BACE1 also resulted in disruption of courtship memory, but not of learning. These defects could be suppressed by administration of a γ-secretase inhibitor.(*232*) BACE1 homozygous mutant mice exhibit memory defects using several behavioral paradigms. These memory deficits can be ameliorated by expressing transgenes encoding APP and Presenillin mutants that result in elevated AICD production.(*233*) As AICD nuclear signaling is largely driven by β-secretase initiated processing of APP, it stands to reason that the LTM specific memory defects resulting from APPL knockdown observed in Drosophila could result from the loss of AICD nuclear signaling.(*185*, *186*, *234*) However, The complexities of APPL processing and resultant signaling preclude drawing such a conclusion based upon existing evidence.

The cognitive defects resulting from AD burden its sufferers and their caretakers in uniquely difficult ways. Though the exploration of AD pathology has yet to lead to effective treatments, it has added to our understanding of the basic mechanisms involved in regulating neural circuits, synaptic plasticity, and memory formation. Substantial evidence implicates aberrant APP processing as causal in AD. However, the role of proteolytic processing of APP and its homologues in normal neuronal function is poorly understood. Recent work indicates that β-secretase initiated processing of APP family proteins is vital for normal brain function, and not solely a disease causing event. In fact, mounting evidence suggest that it is a

consequence of normal neuronal activity. Research in *Drosophila* has contributed to our knowledge of these systems, and with the recent discovery of dBACE and dAβ, research in flies is poised to make growing contributions. Behavioral genetics studies in *Drosophila* unveiled detailed aspects of memory formation that were previously unknown. Recently, APPL was shown to be required specifically for long term memory formation. In chapter III, I will discuss results of our work, indicating that dBACE is also required specifically for LTM, is upregulated as a result of an aversive olfactory conditioning paradigm that produces long term memory, and that this upregulation results in increased AICD production.

Literature Cited

1. E. R. Kandel, *In Search of Memory: The Emergence of a New Science of Mind* (W. W. Norton & Company, 2007).

2. T. V. Bliss, T. Lomo, Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol* **232**, 331–356 (1973).

3. M. A. Sutton, S. E. Masters, M. W. Bagnall, T. J. Carew, Molecular Mechanisms Underlying a Unique Intermediate Phase of Memory inAplysia. *Neuron* **31**, 143–154 (2001).

4. C. H. Bailey, E. R. Kandel, in Progress in Brain Research. (Elsevier, 2008), vol. 169, pp. 179–198.

5. T. Tully, Drosophila learning: behavior and biochemistry. *Behav. Genet.* **14**, 527–557 (1984).

6. A. C. Keene, S. Waddell, Drosophila olfactory memory: single genes to

complex neural circuits. *Nature Reviews Neuroscience* **8**, 341–354 (2007).

7. R. L. Davis, Traces of Drosophila memory. *Neuron* **70**, 8–19 (2011).

8. E. R. Kandel, J. Schwartz, T. Jessell, *Principles of Neural Science, Fifth Edition* (McGraw Hill Professional, 2013).

9. M. Mayford, S. A. Siegelbaum, E. R. Kandel, Synapses and Memory Storage. *Cold Spring Harb Perspect Biol* **4**, a005751–a005751 (2012).

10. A. H. Brand, N. Perrimon, Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).

11. B. D. Pfeiffer *et al.*, Tools for neuroanatomy and neurogenetics in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9715–9720 (2008).

12. V. Ruta *et al.*, A dimorphic pheromone circuit in Drosophila from sensory input to descending output. *Nature* **468**, 686–690 (2010).

13. L. Kahsai, T. Zars, in International Review of Neurobiology. (Elsevier, 2011), vol. 99, pp. 139–167.

14. J. W. Wang, A. M. Wong, J. Flores, L. B. Vosshall, R. Axel, Two-Photon Calcium Imaging Reveals an Odor-Evoked Map of Activity in the Fly Brain. *Cell* **112**, 271–282 (2003).

15. A. Claridge-Chang *et al.*, Writing Memories with Light-Addressable Reinforcement Circuitry. *Cell* **139**, 405–415 (2009).

16. C.-L. Wu, A.-S. Chiang, Genes and circuits for olfactory-associated long-term memory in Drosophila. *J. Neurogenet.* **22**, 257–284 (2008).

17. G. Liu *et al.*, Distinct memory traces for two visual features in the Drosophila brain. *Nature* **439**, 551–556 (2006).

18. C. Schnaitmann, K. Vogt, T. Triphan, H. Tanimoto, Appetitive and aversive visual learning in freely moving Drosophila. *Front Behav Neurosci* **4**, 10–10 (2010).

19. J. Foucaud, J. G. Burns, F. Mery, T. Zars, Ed. Use of spatial information and search strategies in a water maze analog in Drosophila melanogaster. *PLoS ONE* **5**, e15231 (2010).

20. P. Masek, K. Scott, Limited taste discrimination in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14833–14838 (2010).

21. L. C. Griffith, A. Ejima, Courtship learning in Drosophila melanogaster:

diverse plasticity of a reproductive behavior. *Learn Mem* **16**, 743–750 (2009).

22. M.-C. Wu *et al.*, Optogenetic control of selective neural activity in multiple freely moving Drosophila adults. *PNAS* **111**, 5367–5372 (2014).

23. A. Couto, M. Alenius, B. J. Dickson, Molecular, Anatomical, and Functional Organization of the Drosophila Olfactory System. *Current Biology* **15**, 1535–1547 (2005).

24. S. Murakami *et al.*, OptimizingDrosophilaolfactory learning with a semi-automated training device. *J Neurosci Methods* **188**, 195–204 (2010).

25. W. G. Quinn, P. P. Sziber, R. Booker, The Drosophila memory mutant amnesiac. *Nature* **277**, 212–214 (1979).

26. Y. Dudai, Y. N. Jan, D. Byers, W. G. Quinn, S. Benzer, dunce, a mutant of Drosophila deficient in learning. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1684–1688 (1976).

27. K. W. Choi, R. F. Smith, R. M. Buratowski, W. G. Quinn, Deficient protein kinase C activity in turnip, a Drosophila learning mutant. *Journal of Biological Chemistry* **266**, 15999–15606 (1991).

28. E. Perisse, C. Burke, W. Huetteroth, S. Waddell, Shocking Revelations and Saccharin Sweetness Review in the Study of Drosophila Olfactory Memory. *CURBIO* **23**, R752–R763 (2013).

29. C. Margulies, T. Tully, J. Dubnau, Deconstructing memory in Drosophila. *CURBIO* **15**, R700–13 (2005).

30. M. B. Feany, W. G. Quinn, A neuropeptide gene defined by the Drosophila memory mutant amnesiac. *Science* **268**, 869–873 (1995).

31. H. LaFerriere, K. Speichinger, A. Stromhaug, T. Zars, A. Samuel, Ed. The Radish Gene Reveals a Memory Component with Variable Temporal Properties. *PLoS ONE* **6**, e24557 (2011).

32. P.-T. Lee *et al.*, Serotonin-mushroom body circuit modulating the formation of anesthesia-resistant memory in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13794–13799 (2011).

33. V. Goguel *et al.*, Drosophila amyloid precursor protein-like is required for long-term memory. *The Journal of Neuroscience* **31**, 1032–1037 (2011).

34. J. Dubnau, A.-S. Chiang, Systems memory consolidation in Drosophila. *Curr Opin Neurobiol* **23**, 84–91 (2013).

35. Y.-C. Kim, H.-G. Lee, K.-A. Han, D1 dopamine receptor dDA1 is required in the mushroom body neurons for aversive and appetitive learning in Drosophila. *The Journal of Neuroscience* **27**, 7640–7647 (2007).

36. L. B. Vosshall, A. M. Wong, R. Axel, An Olfactory Sensory Map in the Fly Brain. *Cell* **102**, 147–159 (2000).

37. R. L. Davis, Olfactory memory formation in Drosophila: from molecular to systems neuroscience. *Neuroscience* **28**, 275–302 (2005).

38. G. Heimbeck, V. Bugnon, N. Gendre, A. Keller, R. F. Stocker, A central neural circuit for experience-independent olfactory and courtship behavior in Drosophila melanogaster. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 15336–15341 (2001).

39. Z. Lei, K. Chen, H. Li, H. Liu, A. Guo, The GABA system regulates the sparse coding of odors in the mushroom bodies of Drosophila. *Biochem. Biophys. Res. Commun.* **436**, 35–40 (2013).

40. Y. Aso *et al.*, E. Rulifson, Ed. Three dopamine pathways induce aversive odor memories with different stability. *PLoS Genet.* **8**, e1002768–e1002768 (2012).

41. Y. Aso *et al.*, Specific Dopaminergic Neurons for the Formation of Labile Aversive Memory. *Current Biology* **20**, 1445–1451 (2010).

42. Y. Aso *et al.*, Specific dopaminergic neurons for the formation of labile aversive memory. *Curr. Biol.* **20**, 1445–1451 (2010).

43. H. Qin *et al.*, Gamma neurons mediate dopaminergic input during aversive olfactory memory formation in Drosophila. *Curr. Biol.* **22**, 608–614 (2012).

44. C. J. Burke *et al.*, Layered reward signalling through octopamine and dopamine in Drosophila. *Nature* **492**, 433–437 (2012).

45. C. Liu *et al.*, A subset of dopamine neurons signals reward for odour memory in Drosophila. *Nature* **488**, 512–516 (2012).

46. C.-L. Wu *et al.*, Heterotypic gap junctions between two neurons in the drosophila brain are critical for memory. *Curr. Biol.* **21**, 848–854 (2011).

47. J. L. Pitman *et al.*, A pair of inhibitory neurons are required to sustain labile memory in the Drosophila mushroom body. *Curr. Biol.* **21**, 855–861 (2011).

48. J. T. Dimos *et al.*, Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* **321**, 1218–1221 (2008).

49. T. R. Cech, J. A. Steitz, The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. *Cell* **157**, 77–94 (2014).

50. P. Svoboda, Renaissance of mammalian endogenous RNAi. *FEBS Letters* (2014), doi:10.1016/j.febslet.2014.05.030.

51. S. E. Castel, R. A. Martienssen, RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature Reviews Genetics* **14**, 100–112 (2013).

52. G. Meister, Argonaute proteins: functional insights and emerging roles. *Nature Reviews Genetics* **14**, 447–459 (2013).

53. D. Fagegaltier *et al.*, The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. *PNAS* **106**, 21258–21263 (2009).

54. F. M. Cernilogar *et al.*, Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. *Nature* **480**, 391–395 (2011).

55. J. M. Taliaferro *et al.*, Two new and distinct roles for Drosophila Argonaute-2 in the nucleus: alternative pre-mRNA splicing and transcriptional repression. *Genes \& Development* **27**, 378–389 (2013).

56. C. D. Malone, G. J. Hannon, Small RNAs as guardians of the genome. *Cell* **136**, 656–668 (2009).

57. D. Moazed, Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**, 413–420 (2009).

58. C. Li *et al.*, Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* **137**, 509–521 (2009).

59. P. N. Perrat *et al.*, Transposition-driven genomic heterogeneity in the Drosophila brain. *Science* **340**, 91–95 (2013).

60. P. Rajasethupathy *et al.*, A Role for Neuronal piRNAs in the Epigenetic Control of Memory-Related Synaptic Plasticity. *Cell* **149**, 693–707 (2012).

61. Z. Yan *et al.*, Widespread expression of piRNA-like molecules in somatic tissues. *Nucleic Acids Research* **39**, 6596–6607 (2011).

62. A. Janic, L. Mendizabal, S. Llamazares, D. Rossell, C. Gonzalez, Ectopic expression of germline genes drives malignant brain tumor growth in Drosophila. *Science* **330**, 1824–1827 (2010).

63. B. Czech, G. J. Hannon, Small RNA sorting: matchmaking for Argonautes. *Nature Reviews Genetics* **12**, 19–31 (2011).

64. V. N. Kim, J. Han, M. C. Siomi, Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **10**, 126–139 (2009).

65. K. Miyoshi, T. Miyoshi, J. V. Hartig, H. Siomi, M. Siomi, Molecular mechanisms that funnel RNA precursors into endogenous small-interfering RNA and microRNA biogenesis pathways in Drosophila. *RNA (New York, N.Y.)* **16**, 506–515 (2010).

66. P.-H. Wu, M. Isaji, R. W. Carthew, Functionally Diverse MicroRNA Effector Complexes Are Regulated by Extracellular Signaling. *Mol. Cell* **52**, 113–123 (2013).

67. C. Pritchard, H. Cheng, M. Tewari, MicroRNA profiling: approaches and considerations. *Nature Reviews Genetics* **13**, 358–369 (2012).

68. I. P. Sudhakaran *et al.*, FMRP and Ataxin-2 function together in long-term olfactory habituation and neuronal translational control. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E99–E108 (2014).

69. X. Li, P. Jin, Roles of small regulatory RNAs in determining neuronal identity. *Nat. Rev. Neurosci.* **11**, 329–338 (2010).

70. K. Tsurudome *et al.*, The Drosophila miR-310 Cluster Negatively Regulates Synaptic Strength at the Neuromuscular Junction. *Neuron* **68**, 879–893 (2010).

71. N. R. Smalheiser, G. Lugli, J. Thimmapuram, E. H. Cook, J. Larson, Endogenous siRNAs and noncoding RNA-derived small RNAs are expressed in adult mouse hippocampus and are up-regulated in olfactory discrimination training. *RNA* **17**, 166–181 (2011).

72. B.-T. Juang *et al.*, Endogenous nuclear RNAi mediates behavioral adaptation to odor. *Cell* **154**, 1010–1022 (2013).

73. Z. Yu, X. Teng, N. M. Bonini, C. E. Pearson, Ed. Triplet Repeat–Derived siRNAs Enhance RNA–Mediated Toxicity in a Drosophila Model for Myotonic Dystrophy. *PLoS Genet.* **7**, e1001340 (2011).

74. R. A. Martienssen, E. J. Richards, DNA methylation in eukaryotes. *Curr. Opin. Genet. Dev.* **5**, 234–242 (1995).

75. M. A. Matzke, M. Primig, J. Trnovsky, A. J. Matzke, Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. *The EMBO Journal* **8**, 643–649 (1989).

76. Y. D. Park *et al.*, Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in

methylation and gene activity. *The Plant Journal* **9**, 183–194 (1996).

77. M. Wassenegger, S. Heimes, L. Riedel, H. L. Sänger, RNA-directed de novo methylation of genomic sequences in plants. *Cell* **76**, 567–576 (1994).

78. J. A. Lindbo, L. Silva-Rosales, W. M. Proebsting, W. G. Dougherty, Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. *The Plant Cell* **5**, 1749 (1993).

79. M. A. Matzke, A. J. M. Matzke, Planting the Seeds of a New Paradigm. *PLoS Biol* **2**, e133 (2004).

80. A. Fire *et al.*, Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806–811 (1998).

81. P. D. Zamore, T. Tuschl, P. A. Sharp, D. P. Bartel, RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33 (2000).

82. S. M. Elbashir, W. Lendeckel, T. Tuschl, RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes \& Development* **15**, 188–200 (2001).

83. S. M. Hammond, E. Bernstein, D. Beach, G. J. Hannon, An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293–296 (2000).

84. E. Bernstein, A. A. Caudy, S. M. Hammond, G. J. Hannon, Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).

85. K. M. Nishida *et al.*, Roles of R2D2, a cytoplasmic D2 body component, in the endogenous siRNA pathway in Drosophila. *Molecular Cell* **49**, 680–691 (2013).

86. J. T. Marques *et al.*, Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila. *Nat. Struct. Mol. Biol.* **17**, 24–30 (2010).

87. C. Matranga, Y. Tomari, C. Shin, D. P. Bartel, P. D. Zamore, Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123**, 607–620 (2005).

88. S. Iwasaki *et al.*, Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. *Mol. Cell* **39**, 292–299 (2010).

89. Y. Liu *et al.*, C3PO, an endoribonuclease that promotes RNAi by facilitating

RISC activation. *Science* **325**, 750–753 (2009).

90. M. D. Horwich *et al.*, The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *CURBIO* **17**, 1265–1272 (2007).

91. S. L. Ameres, J.-H. Hung, J. Xu, Z. Weng, P. D. Zamore, Target RNA-directed tailing and trimming purifies the sorting of endo-siRNAs between the two Drosophila Argonaute proteins. *RNA (New York, N.Y.)* **17**, 54–63 (2011).

92. R. P. van Rij *et al.*, The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. *Genes \& Development* **20**, 2985–2995 (2006).

93. D. Galiana-Arnoux, C. Dostert, A. Schneemann, J. A. Hoffmann, J.-L. Imler, Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila. *Nature Immunology* **7**, 590–597 (2006).

94. B. Czech *et al.*, An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**, 798–802 (2008).

95. Y. Kawamura *et al.*, Drosophila endogenous small RNAs bind to Argonaute|[thinsp]|2 in somatic cells. *Nature* **453**, 793–797 (2008).

96. K. Okamura, E. Lai, Endogenous small interfering RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **9**, 673–678 (2008).

97. K. Okamura *et al.*, The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**, 803–806 (2008).

98. D. Golden, V. Gerbasi, E. Sontheimer, An Inside Job for siRNAs. *Molecular Cell* **31**, 309–312 (2008).

99. M. Ghildiyal *et al.*, Endogenous siRNAs Derived from Transposons and mRNAs in Drosophila Somatic Cells. *Science* **320**, 1077–1081 (2008).

100. W.-J. Chung, K. Okamura, R. Martin, E. Lai, Endogenous RNA Interference Provides a Somatic Defense against Drosophila Transposons. *Current Biology* **18**, 795–802 (2008).

101. N. Moshkovich *et al.*, RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function. *Genes \& Development* **25**, 1686–1701 (2011).

102. E. G. Moss, L. Tang, Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Developmental biology* **258**, 432–442 (2003).

103. E. C. Lai, microRNAs: Runts of the Genome Assert Themselves. *CURBIO* **13**,

R925–R936 (2003).

104. R. C. Lee, R. L. Feinbaum, V. Ambros, The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–854 (1993).

105. B. Wightman, I. Ha, G. Ruvkun, Posttranscriptional regulation of the heterochronic genelin-14bylin-4mediates temporal pattern formation in C. elegans. *Cell* **75**, 855–862 (1993).

106. J.-S. Yang, E. C. Lai, Alternative miRNA Biogenesis Pathways and the Interpretation of Core miRNA Pathway Mutants. *Molecular Cell* **43**, 892–903 (2011).

107. S. L. Ameres, P. D. Zamore, Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–488 (2013).

108. R. Fukunaga *et al.*, Dicer partner proteins tune the length of mature miRNAs in flies and mammals. *Cell* **151**, 533–546 (2012).

109. M. Ghildiyal, J. Xu, H. Seitz, Z. Weng, P. D. Zamore, Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. *RNA (New York, N.Y.)* **16**, 43–56 (2010).

110. S. L. Ameres *et al.*, Target RNA-directed trimming and tailing of small silencing RNAs. *Science* **328**, 1534–1539 (2010).

111. R. J. Taft *et al.*, Small RNAs derived from snoRNAs. *RNA (New York, N.Y.)* **15**, 1233–1240 (2009).

112. C. Ender *et al.*, A human snoRNA with microRNA-like functions. *Mol. Cell* **32**, 519–528 (2008).

113. modENCODE Consortium *et al.*, Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* **330**, 1787–1797 (2010).

114. E. Berezikov *et al.*, Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. *Genome research* **21**, 203–215 (2011).

115. J. G. Ruby, C. H. Jan, D. P. Bartel, Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**, 83–86 (2007).

116. K. Okamura, J. W. Hagen, H. Duan, D. M. Tyler, E. C. Lai, The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell* **130**, 89–100 (2007).

117. A. Flynt, J. Greimann, W.-J. Chung, C. Lima, E. Lai, MicroRNA Biogenesis via Splicing and Exosome-Mediated Trimming in Drosophila. *Molecular Cell* **38**, 900–907 (2010).

118. E. A. Glazov *et al.*, A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome research* **18**, 957–964 (2008).

119. E. Berezikov, W.-J. Chung, J. Willis, E. Cuppen, E. C. Lai, Mammalian mirtron genes. *Molecular Cell* **28**, 328–336 (2007).

120. J. E. Babiarz, J. G. Ruby, Y. Wang, D. P. Bartel, R. Blelloch, Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes \& Development* **22**, 2773–2785 (2008).

121. V. Auyeung, I. Ulitsky, S. McGeary, D. Bartel, Beyond Secondary Structure: Primary-Sequence Determinants License Pri-miRNA Hairpins for Processing. *Cell* **152**, 844–858 (2013).

122. T. Nishihara, L. Zekri, J. E. Braun, E. Izaurralde, miRISC recruits decapping factors to miRNA targets to enhance their degradation. *Nucleic Acids Research* **41**, 8692–8705 (2013).

123. I. Behm-Ansmant *et al.*, mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes \& Development* **20**, 1885–1898 (2006).

124. T. Fukaya, Y. Tomari, MicroRNAs Mediate Gene Silencing via Multiple Different Pathways in Drosophila. *Molecular Cell* **48**, 825–836 (2012).

125. R. Thermann, M. W. Hentze, Drosophila miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature* **447**, 875–878 (2007).

126. S. Banerjee, P. Neveu, K. S. Kosik, A coordinated local translational control point at the synapse involving relief from silencing and MOV10 degradation. *Neuron* **64**, 871–884 (2009).

127. K. Miyoshi, T. N. Okada, H. Siomi, M. C. Siomi, Characterization of the miRNA-RISC loading complex and miRNA-RISC formed in the Drosophila miRNA pathway. *RNA (New York, N.Y.)* **15**, 1282–1291 (2009).

128. R. J. Ross, M. M. Weiner, H. Lin, PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* **505**, 353–359 (2014).

129. P. M. Guzzardo, F. Muerdter, G. J. Hannon, The piRNA pathway in flies: highlights and future directions. *Curr. Opin. Genet. Dev.* **23**, 44–52 (2013).

130. J. Brennecke *et al.*, Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. *Cell* **128**, 1089–1103 (2007).

131. K. Saito *et al.*, A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. *Nature* **461**, 1296–1299 (2009).

132. N. Robine *et al.*, A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr. Biol.* **19**, 2066–2076 (2009).

133. H. Nishimasu *et al.*, Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* **491**, 284–287 (2012).

134. F. Voigt *et al.*, Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc. *RNA (New York, N.Y.)* **18**, 2128–2134 (2012).

135. S. Kawaoka, N. Izumi, S. Katsuma, Y. Tomari, 3' end formation of PIWI-interacting RNAs in vitro. *Mol. Cell* **43**, 1015–1022 (2011).

136. J. Brennecke *et al.*, An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing. *Science* **322**, 1387–1392 (2008).

137. C. D. Malone *et al.*, Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **137**, 522–535 (2009).

138. A. de Vanssay *et al.*, Paramutation in Drosophila linked to emergence of a piRNA-producing locus. *Nature* **490**, 112–115 (2012).

139. A. Nagao *et al.*, Biogenesis pathways of piRNAs loaded onto AGO3 in the Drosophila testis. *RNA (New York, N.Y.)* **16**, 2503–2515 (2010).

140. L. S. Gunawardane *et al.*, A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science* **315**, 1587–1590 (2007).

141. D. Handler *et al.*, A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *The EMBO Journal* **30**, 3977–3993 (2011).

142. M. S. Klenov *et al.*, Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *PNAS* **108**, 18760–18765 (2011).

143. K. Saito *et al.*, Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. *Genes \& Development* **24**, 2493–2498 (2010).

144. N. Darricarrère, N. Liu, T. Watanabe, H. Lin, Function of Piwi, a nuclear Piwi/Argonaute protein, is independent of its slicer activity. *PNAS* **110**, 1297–1302 (2013).

145. G. Sienski, D. Dönertas, J. Brennecke, Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**, 964–980 (2012).

146. A. Le Thomas *et al.*, Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes \& Development* **27**, 390–399 (2013).

147. N. V. Rozhkov, M. Hammell, G. J. Hannon, Multiple roles for Piwi in silencing Drosophila transposons. *Genes \& Development* **27**, 400–412 (2013).

148. V. K. Gangaraju *et al.*, Drosophila Piwi functions in Hsp90-mediated suppression of phenotypic variation. *Nat. Genet.* **43**, 153–158 (2011).

149. V. Specchia *et al.*, Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. *Nature* **463**, 662–665 (2010).

150. X. Huang *et al.*, A major epigenetic programming mechanism guided by piRNAs. *Developmental Cell* **24**, 502–516 (2013).

151. X. Ma *et al.*, C. Bökel, Ed. Piwi is required in multiple cell types to control germline stem cell lineage development in the Drosophila ovary. *PLoS ONE* **9**, e90267 (2014).

152. K. Meier *et al.*, A. Akhtar, Ed. LINT, a Novel dL(3)mbt-Containing Complex, Represses Malignant Brain Tumour Signature Genes. *PLoS Genet.* **8**, e1002676 (2012).

153. R. Darnell, RNA Protein Interaction in Neurons. *Annu. Rev. Neurosci.* **36**, 243–270 (2013).

154. J. Morante, D. M. Vallejo, C. Desplan, M. Dominguez, Conserved miR-8/miR-200 defines a glial niche that controls neuroepithelial expansion and neuroblast transition. *Developmental Cell* **27**, 174–187 (2013).

155. M. Kucherenko, J. Barth, A. Fiala, H. Shcherbata, Steroid-induced microRNA let-7 acts as a spatio-temporal code for neuronal cell fate in the developing Drosophila brain. *The EMBO Journal* **advance online publication** (2012), doi:10.1038/emboj.2012.298.

156. Y.-C. Wu, C.-H. Chen, A. Mercer, N. S. Sokol, let-7-Complex MicroRNAs Regulate the Temporal Identity of Drosophila Mushroom Body Neurons via chinmo. *Developmental Cell* **23**, 202–209 (2012).

157. T. M. Huckaba, A. Gennerich, J. E. Wilhelm, A. H. Chishti, R. D. Vale, Kinesin-73 is a processive motor that localizes to Rab5-containing organelles. *J. Biol. Chem.* **286**, 7457–7467 (2011).

158. P. Kner, B. B. Chhun, E. R. Griffis, L. Winoto, M. G. L. Gustafsson, Super-resolution video microscopy of live cells by structured illumination. *Nat. Methods* **6**, 339–342 (2009).

159. D. A. Wagh *et al.*, Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in Drosophila. *Neuron* **49**, 833–844 (2006).

160. W. Li *et al.*, MicroRNA-276a functions in ellipsoid body and mushroom body neurons for naive and conditioned olfactory avoidance in Drosophila. *The Journal of Neuroscience* **33**, 5821–5833 (2013).

161. P. Rajasethupathy *et al.*, Characterization of Small RNAs in Aplysia Reveals a Role for miR-124 in Constraining Synaptic Plasticity through CREB. *Neuron* **63**, 803–817 (2009).

162. K. R. Nesler *et al.*, T. H. Gillingwater, Ed. The miRNA Pathway Controls Rapid Changes in Activity-Dependent Synaptic Structure at the Drosophila melanogaster Neuromuscular Junction. *PLoS ONE* **8**, e68385 (2013).

163. M. G. Thomas, M. L. Pascual, D. Maschi, L. Luchelli, G. L. Boccaccio, Synaptic control of local translation: the plot thickens with new characters. *Cell. Mol. Life Sci.* **71**, 2219–2239 (2014).

164. I. J. Cajigas *et al.*, The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* **74**, 453–466 (2012).

165. L. F. Gumy *et al.*, Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. *RNA (New York, N.Y.)* **17**, 85–98 (2011).

166. T. Rogerson *et al.*, Synaptic tagging during memory allocation. *Nature Reviews Neuroscience* **15**, 157–169 (2014).

167. E. McNeill, D. Van Vactor, MicroRNAs Shape the Neuronal Landscape. *Neuron* **75**, 363–379 (2012).

168. B. Pai *et al.*, NMDA receptor-dependent regulation of miRNA expression and association with Argonaute during LTP in vivo. *Front. Cell. Neurosci.* **7** (2014), doi:10.3389/fncel.2013.00285.

169. S. Bicker *et al.*, The DEAH-box helicase DHX36 mediates dendritic localization of the neuronal precursor-microRNA-134. *Genes \& Development* **27**, 991–996 (2013).

170. G. M. Schratt *et al.*, A brain-specific microRNA regulates dendritic spine

development. *Nature* **439**, 283–289 (2006).

171. S. I. Ashraf, A. L. McLoon, S. M. Sclarsic, S. Kunes, Synaptic Protein Synthesis Associated with Memory Is Regulated by the RISC Pathway in Drosophila. *Cell* **124**, 191–205 (2006).

172. T. J. Jarome, C. T. Werner, J. L. Kwapis, F. J. Helmstetter, Activity dependent protein degradation is critical for the formation and stability of fear memory in the amygdala. *PLoS ONE* **6**, e24349 (2011).

173. L. H. Gregersen *et al.*, MOV10 Is a 5′ to 3′ RNA Helicase Contributing to UPF1 mRNA Target Degradation by Translocation along 3′ UTRs. *Molecular Cell* **54**, 573–585 (2014).

174. N. M. Sosanya *et al.*, Degradation of high affinity HuD targets releases Kv1.1 mRNA from miR-129 repression by mTORC1. *The Journal of Cell Biology* **202**, 53–69 (2013).

175. P. Kundu, M. R. Fabian, N. Sonenberg, S. N. Bhattacharyya, W. Filipowicz, HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA. *Nucleic Acids Research* **40**, 5088–5100 (2012).

176. U. Bräuer, E. Zaharieva, M. Soller, Regulation of ELAV/Hu RNA-binding proteins by phosphorylation. *Biochem. Soc. Trans.* **42**, 1147–1151 (2014).

177. A. Doller, J. Pfeilschifter, W. Eberhardt, Signalling pathways regulating nucleo-cytoplasmic shuttling of the mRNA-binding protein HuR. *Cellular Signalling* **20**, 2165–2173 (2008).

178. R. Muddashetty *et al.*, Reversible Inhibition of PSD-95 mRNA Translation by miR-125a, FMRP Phosphorylation, and mGluR Signaling. *Molecular Cell* **42**, 673–688 (2011).

179. C.-K. Park *et al.*, Extracellular MicroRNAs Activate Nociceptor Neurons to Elicit Pain via TLR7 and TRPA1. *Neuron* **82**, 47–54 (2014).

180. Z. Lv *et al.*, G. Camussi, Ed. Argonaute 2 in Cell-Secreted Microvesicles Guides the Function of Secreted miRNAs in Recipient Cells. *PLoS ONE* **9**, e103599 (2014).

181. B. J. Goldie *et al.*, Activity-associated miRNA are packaged in Map1b-enriched exosomes released from depolarized neurons. *Nucleic Acids Research* **42**, 9195–9208 (2014).

182. J. Wen *et al.*, Diversity of miRNAs, siRNAs, and piRNAs across 25 Drosophila cell lines. *Genome research* **24**, 1236–1250 (2014).

183. N. D. L'Etoile, C. I. Bargmann, Olfaction and odor discrimination are mediated by the C. elegans guanylyl cyclase ODR-1. *Neuron* **25**, 575–586 (2000).

184. J. Kang *et al.*, The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature* **325**, 733–736 (1987).

185. N. D. Belyaev *et al.*, The transcriptionally active amyloid precursor protein (APP) intracellular domain is preferentially produced from the 695 isoform of APP in a {beta}-secretase-dependent pathway. *J. Biol. Chem.* **285**, 41443–41454 (2010).

186. Z. V. Goodger *et al.*, Nuclear signaling by the APP intracellular domain occurs predominantly through the amyloidogenic processing pathway. *J. Cell. Sci.* **122**, 3703–3714 (2009).

187. S. E. Hoey *et al.*, A. I. Bush, Ed. AMPA receptor activation promotes non-amyloidogenic amyloid precursor protein processing and suppresses neuronal amyloid-β production. *PLoS ONE* **8**, e78155 (2013).

188. S. E. Hoey, R. J. Williams, M. S. Perkinton, Synaptic NMDA receptor activation stimulates alpha-secretase amyloid precursor protein processing and inhibits amyloid-beta production. *The Journal of Neuroscience* **29**, 4442–4460 (2009).

189. J. O. Rinne *et al.*, 11C-PiB PET assessment of change in fibrillar amyloid-β load in patients with Alzheimer's disease treated with bapineuzumab: a phase 2, double-blind, placebo-controlled, ascending-dose study. *The Lancet Neurology* **9**, 363–372 (2010).

190. D. Puzzo *et al.*, Endogenous amyloid-β is necessary for hippocampal synaptic plasticity and memory. *Annals of Neurology* **69**, 819–830 (2011).

191. L. Luo, L. E. Martin-Morris, K. White, Identification, secretion, and neural expression of APPL, a Drosophila protein similar to human amyloid protein precursor. *J. Neurosci.* **10**, 3849–3861 (1990).

192. L. Luo, T. Tully, K. White, Human amyloid precursor protein ameliorates behavioral deficit of flies deleted forapplgene. *Neuron* **9**, 595–605 (1992).

193. I. Greeve *et al.*, Age-dependent neurodegeneration and Alzheimer-amyloid plaque formation in transgenic Drosophila. *The Journal of Neuroscience* **24**, 3899–3906 (2004).

194. K. Carmine-Simmen *et al.*, Neurotoxic effects induced by the Drosophila amyloid-beta peptide suggest a conserved toxic function. *Neurobiology of Disease* **33**, 274–281 (2009).

195. C. Groth, W. G. Alvord, O. A. Quinones, M. E. Fortini, Pharmacological analysis of Drosophila melanogaster gamma-secretase with respect to differential proteolysis of Notch and APP. *Molecular Pharmacology* **77**, 567–574 (2010).

196. K. Iijima *et al.*, Dissecting the pathological effects of human Abeta40 and Abeta42 in Drosophila: a potential model for Alzheimer's disease. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6623–6628 (2004).

197. B. Poeck, R. Strauss, D. Kretzschmar, Analysis of amyloid precursor protein function in Drosophila melanogaster. *Exp Brain Res* **217**, 413–421 (2012).

198. K. Prüßing, A. Voigt, J. B. Schulz, Drosophila melanogaster as a model organism for Alzheimer's disease. *Molecular Neurodegeneration* **8**, 35 (2013).

199. B. E. Needham *et al.*, Identification of the Alzheimer's disease amyloid precursor protein (APP) and its homologue APLP2 as essential modulators of glucose and insulin homeostasis and growth. *J. Pathol.* **215**, 155–163 (2008).

200. K. T. Jacobsen, K. Iverfeldt, Amyloid precursor protein and its homologues: a family of proteolysis-dependent receptors. *Cell. Mol. Life Sci.* **66**, 2299–2318 (2009).

201. N. N. Nalivaeva, A. J. Turner, The amyloid precursor protein: A biochemical enigma in brain development, function and disease. *FEBS Letters* **587**, 2046–2054 (2013).

202. A. Marchler-Bauer, S. H. Bryant, CD-Search: protein domain annotations on the fly. *Nucleic Acids Research* **32**, W327–31 (2004).

203. A. Marchler-Bauer *et al.*, CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Research* **41**, D348–52 (2013).

204. L. Torroja, M. Packard, M. Gorczyca, K. White, V. Budnik, The Drosophila beta-amyloid precursor protein homolog promotes synapse differentiation at the neuromuscular junction. *The Journal of Neuroscience* **19**, 7793–7803 (1999).

205. J. S. Wentzell *et al.*, Amyloid precursor proteins are protective in Drosophila models of progressive neurodegeneration. *Neurobiology of Disease* **46**, 78–87 (2012).

206. U. Das *et al.*, Activity-Induced Convergence of APP and BACE-1 in Acidic Microdomains via an Endocytosis-Dependent Pathway. *Neuron* **79**, 447–460 (2013).

207. A. D. Haase *et al.*, Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. *Genes \& Development* **24**, 2499–2504 (2010).

208. L. Torroja, L. Luo, K. White, APPL, the Drosophila member of the APP-family, exhibits differential trafficking and processing in CNS neurons. *J. Neurosci.* **16**, 4638–4650 (1996).

209. Y. Li, T. Liu, Y. Peng, C. Yuan, A. Guo, Specific functions of Drosophila amyloid precursor-like protein in the development of nervous system and nonneural tissues. *J. Neurobiol.* **61**, 343–358 (2004).

210. D. A. S. S. H. S. R. B. D. S. B. A. H. Maarten Leyssen, Amyloid precursor protein promotes post-developmental neurite arborization in the Drosophila brain. *The EMBO Journal* **24**, 2944–2955 (2005).

211. K. Furukawa, S. W. Barger, E. M. Blalock, M. P. Mattson, Activation of K+ channels and suppression of neuronal activity by secreted beta-amyloid-precursor protein. *Nature* **379**, 74–78 (1996).

212. K. Furukawa, M. P. Mattson, Secreted amyloid precursor protein α selectively suppresses N-methyl-d-aspartate currents in hippocampal neurons: involvement of cyclic GMP. *Neuroscience* **83**, 429–438 (1998).

213. Y. Wang, G. Zhang, H. Zhou, A. Barakat, H. Querfurth, K. Hashimoto, Ed. Opposite Effects of Low and High Doses of Aβ42 on Electrical Network and Neuronal Excitability in the Rat Prefrontal Cortex. *PLoS ONE* **4**, e8366 (2009).

214. D. L. Pettit, Z. Shao, J. L. Yakel, beta-Amyloid(1-42) peptide directly modulates nicotinic receptors in the rat hippocampal slice. *The Journal of Neuroscience* **21**, RC120 (2001).

215. M. Ohno *et al.*, BACE1 Deficiency Rescues Memory Deficits and Cholinergic Dysfunction in a Mouse Model of Alzheimer's Disease. *Neuron* **41**, 27–33 (2004).

216. J. H. Jhamandas *et al.*, Cellular mechanisms for amyloid beta-protein activation of rat cholinergic basal forebrain neurons. *J. Neurophysiol.* **86**, 1312–1320 (2001).

217. F. J. Ekinci, K. U. Malik, T. B. Shea, Activation of the L voltage-sensitive calcium channel by mitogen-activated protein (MAP) kinase following exposure of neuronal cells to beta-amyloid. MAP kinase mediates beta-amyloid-induced neurodegeneration. *Journal of Biological Chemistry* **274**, 30322–30327 (1999).

218. C. R. Raymond, D. R. Ireland, W. C. Abraham, NMDA receptor regulation by amyloid-β does not account for its inhibition of LTP in rat hippocampus. *Brain Research* **968**, 263–272 (2003).

219. S. L. Cousins, N. Innocent, F. A. Stephenson, Neto1 associates with the NMDA receptor/amyloid precursor protein complex. *Journal of Neurochemistry* **126**, 554–564 (2013).

220. T. G. Beach, Y. M. Kuo, C. Schwab, D. G. Walker, A. E. Roher, Reduction of cortical amyloid beta levels in guinea pig brain after systemic administration of physostigmine. *Neurosci. Lett.* **310**, 21–24 (2001).

221. R. M. Nitsch, B. E. Slack, R. J. Wurtman, J. H. Growdon, Release of Alzheimer amyloid precursor derivatives stimulated by activation of muscarinic acetylcholine receptors. *Science* **258**, 304–307 (1992).

222. R. M. Nitsch, S. A. Farber, J. H. Growdon, R. J. Wurtman, Release of amyloid beta-protein precursor derivatives by electrical depolarization of rat hippocampal slices. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5191–5193 (1993).

223. C. Hock *et al.*, Treatment with the selective muscarinic m1 agonist talsaclidine decreases cerebrospinal fluid levels of A beta 42 in patients with Alzheimer's disease. *Amyloid* **10**, 1–6 (2003).

224. F. Kamenetz *et al.*, APP processing and synaptic function. *Neuron* **37**, 925–937 (2003).

225. J. R. Cirrito *et al.*, Synaptic activity regulates interstitial fluid amyloid-beta levels in vivo. *Neuron* **48**, 913–922 (2005).

226. S. Lesné *et al.*, NMDA receptor activation inhibits alpha-secretase and promotes neuronal amyloid-beta production. *The Journal of Neuroscience* **25**, 9367–9377 (2005).

227. D. K. Verges, J. L. Restivo, W. D. Goebel, D. M. Holtzman, J. R. Cirrito, Opposing synaptic regulation of amyloid-β metabolism by NMDA receptors in vivo. *The Journal of Neuroscience* **31**, 11328–11337 (2011).

228. C. K. G. T. S. S. Christian Haass, Trafficking and Proteolytic Processing of APP. *Cold Spring Harbor Perspectives in Medicine* **2**, a006270–a006270 (2012).

229. P. R. Turner, K. O'Connor, W. P. Tate, W. C. Abraham, Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. *Progress in Neurobiology* **70**, 1–32 (2003).

230. V. Goguel *et al.*, Drosophila Amyloid Precursor Protein-Like Is Required for Long-Term Memory. *The Journal of Neuroscience* **31**, 1032–1037 (2011).

231. S. M. J. McBride *et al.*, Pharmacological and genetic reversal of age-dependent cognitive deficits attributable to decreased presenilin function. *The Journal of Neuroscience* **30**, 9510–9522 (2010).

232. R. Chakraborty *et al.*, Characterization of a Drosophila Alzheimer's Disease Model: Pharmacological Rescue of Cognitive Defects. *PLoS ONE* **6**, e20799 (2011).

233. F. M. Laird *et al.*, BACE1, a major determinant of selective vulnerability of the brain to amyloid-beta amyloidogenesis, is essential for cognitive, emotional, and synaptic functions. *The Journal of Neuroscience* **25**, 11693–11709 (2005).

234. B. B. Flammang *et al.*, Evidence that the amyloid-β protein precursor intracellular domain, AICD, derives from β-secretase-generated C-terminal fragment. *J Alzheimers Dis* **30**, 145–153 (2012).

Chapter II

Drosophila Olfactory Long Term Memory Formation Alters Short Non-Protein Coding RNA Profiles

Summary

Long term memory (LTM) formation involves changes in the activity of neural circuits brought about by alterations in neuronal structure and synaptic efficacy. Changes in transcription and protein synthesis are required for LTM (reviewed in (*1*, *2*)). Short non-protein coding RNAs (sRNAs) transcriptionally and post-transcriptionally regulate expression of genes, including those involved in synaptic plasticity and LTM formation.(*3*-*5*) Previous studies have identified individual sRNAs that are regulated in response to neural activity.(*6*, *7*) To date, few studies have examined genome wide regulation of any class of sRNA, let alone all sRNAs, during memory formation. In this study, we sought to thoroughly catalog microRNAs and other sRNAs expressed in Drosophila heads, and to describe any changes in their expression during LTM formation. We conducted an extensive analysis of sRNA regulation in response to aversive olfactory classical conditioning of *Drosophila*. Following conditioning, we find that miR-312-3p is downregulated, and that miR-314-3p, miR-956-3p, miR-958-3p, and miR-958-5p are upregulated. Our use of high-throughput sequencing (HTS) technology also allowed us to profile isoMiRs, untemplated nucleotide additions to microRNAs, and instances of RNA editing of microRNAs. Similarly, we were able to catalog endogenous short interfering RNAs (esiRNAs) and piwi interacting RNAs (piRNAs) expressed in the *Drosophila* head, and to identify several loci producing these sRNAs that respond to conditioning. Lastly, by examining changes in sRNA expression resulting from presentation of only the unconditioned stimulus (US) or only the conditioned

stimulus (CS), we show that observed changes in sRNA expression following conditioning are primarily driven by the US.

Introduction

While memories that last only a short time can be formed quickly, and after only a brief presentation of the thing to be remembered, formation of lasting memories is slower, and usually requires practice. Even after practice, formation of lasting memories can, for a time, be disrupted by distracting information. However, if the memory persists beyond this labile period, it becomes largely immune to such distracting information, and can last a lifetime.(*8*) In the initial minutes following learning, the ability to correctly recall fleeting or lasting memories decays rapidly. In the case of a fleeting memory this rapid decay continues until, after a few minutes, the memory is lost. However, though a lasting memory also decays rapidly in the minutes following learning, the memory is not lost, and the decay becomes much more gradual in subsequent hours.(*9*) These behavioral observations reflect fundamental and evolutionarily conserved aspects of the biology underlying memory formation. Initially, formation of fleeting and lasting types of memories involve the same mechanisms. However, formation of lasting memories activates additional mechanisms that stabilize the memory.(*2*)

Memories are encoded in a process that involves alterations in the activity of neural circuits. Such alterations in activity occur through changes in the number, structure, or efficacy of synaptic connections. Initially, these changes occur through the modification of proteins present at existing synapses. Such modifications are not stable, and resembling short-term memory (STM), they decay on the order of

minutes. LTM formation initially also involves modification of proteins already in place at existing synapses. However, LTM formation requires the slower processes of transcription and protein synthesis. Newly synthesized proteins and altered transcriptional programs produce lasting changes in neural excitability, connectivity, and synaptic efficacy or structure.(*2*) Some aspects of gene regulation involved in the formation of LTM formation have been characterized, but it is clear that significant work remains in this area.

Much of what is known of the genetic requirements of memory formation comes from studies of *Drosophila*. Though the *Drosophila* brain is orders of magnitude less complex than mammalian brain, it is still capable of generating a diverse set of behaviors. In addition, *Drosophila* can form lasting memories in a variety of learning paradigms. *Drosophila* memory is readily evaluated through observation of the behavior of individuals or of large groups of flies. Though many methods for training flies and evaluating learning have been developed, aversive olfactory classical conditioning is the most widely used and best understood training paradigm.(*10*) In this paradigm, groups of flies are placed in a chamber containing a grid of electrodes and in which odors can be presented and removed at will. During conditioning, flies receive the unconditioned stimulus (US) in the form of a mild shock to their feet via an electrode grid while being presented with the conditioned stimulus (CS) in the form of an odorant of choice. (*11-13*) Memory in flies trained using this method is typically evaluated using a T maze, in which one arm of the T contains the conditioned odorant, and the other arm of the T contains a

neutral odorant. Memory is scored by observing the fraction of flies present in each arm at the end of the trial. A single pairing of shock and odor produces an associative memory that lasts on the order of minutes. Repeated sessions pairing the odor and shock, with periods of rest between sessions, produces robust associative memory that can last for days.(*9*) In classic work using these techniques, Tully and Quinn identified several genes and neural circuitry required during each phase of LTM formation.(*14-17*) Of note, they showed that the ~5000 neurons of the mushroom body (MB) of the *Drosophila* brain comprise the major site of associative olfactory memory storage (Reviewed in (*18*)).

Previous studies have shown that microRNAs and the multiprotein complexes they associate with (termed RNA induced silencing complexes (RISCs)) regulate the physiology of neurons, synaptic plasticity, and behavior, by controlling protein synthesis.(*19-21*) MicroRNAs are involved in the control of protein expression from mRNAs localized at or near synapses, and therefore may have uniquely important functions in regulating synaptic plasticity and memory.(*4*) Additionally, several microRNAs and RISC associated proteins have been shown to be regulated by neural activity.(*20*, *22*, *23*) Because microRNA sequences specify which mRNAs are targeted for RISC induced silencing, changes in microRNA expression profiles during memory formation result in altered patterns of protein synthesis.(*6*) Complicating matters, many if not most mRNAs are predicted targets of multiple microRNAs.(*24*) Though changes in the expression of individual

microRNAs following learning have been documented, genome wide changes in microRNA expression during memory formation remains poorly studied.

The development and application of high throughput sequencing (HTS) of sRNAs has revealed that a diversity of sRNAs are produced from microRNA precursors. Though in most cases, a single 'canonical' sequence from each of the 5' and 3' arms of the pre-microRNA hairpin represent the overwhelming majority of sequencing reads derived from a given microRNA precursor, examination of 'non-canonical' microRNA reads can provide insight into the biogenesis and regulation of microRNAs. Changes in the proportions of canonical and non-canonical microRNA reads derived from a given pre-microRNA may also have important consequences for the regulatory output of the precursor.(*25*) The biological significance of non-canonical microRNA reads is poorly understood, and largely unknown in the context of synaptic plasticity or memory.

Recently, other classes of sRNAs including endogenous short interfering RNAs (esiRNAs) and piwi interacting RNAs (piRNAs) have been implicated in synaptic plasticity and memory formation in nematodes, mollusks, and mammals.(*3*, *5*, *26*) Similar to microRNAs, esiRNA and piRNA sequences specify the targets for silencing via proteins incorporated in RISCs. In *Drosophila*, esiRNAs and piRNAs both suppress mobilization of retrotransposons.(*27*, *28*) Silencing of retrotransposons via the piRNA pathway is non-uniform in the Drosophila brain, and in the MB in particular. Retrotransposon mobilization leads to genetic diversity

in the *Drosophila* brain, and preferentially affects genes involved in neural physiology.(*29*) esiRNAs are also known to target mRNAs for post-transcriptional silencing in *Drosophila*.(*30*, *31*) To date, no published study has examined regulation of esiRNAs or piRNAs during memory formation in *Drosophila*.

Here, I present the results of profiling genome wide sRNA expression in *Drosophila* heads, and identify changes in sRNA expression during LTM formation resulting from aversive olfactory classical conditioning. We profile sRNA expression in heads of flies similarly treated, but exposed only to the US (shock), or to the CS (odor). Our results indicate that changes in sRNA expression during LTM formation are primarily driven by the US, which involves dopaminergic signaling in the MB.

In section I of this chapter, I present a detailed analysis of microRNA expression in the fly head, and the changes in microRNA profiles during LTM formation. I show that miR-312-3p is significantly downregulated, and that miR-314-3p, miR-956-3p, miR-958-3p, and miR-958-5p are upregulated during LTM formation. Using microRNA target prediction and gene ontology (GO) analyses, we have explored the potential regulatory consequences resulting from the observed changes in the expression of these microRNAs. We document an extensive set of non-canonical microRNA reads, but find that LTM formation does not significantly alter the proportion of microRNA reads that they represent.

In section II of this chapter, I discuss our analysis of esiRNA and piRNA profiles in the Drosophila head, and changes in their expression during LTM formation. I catalog esiRNA producing loci, and identify 6 that are regulated during LTM formation or US exposure, and none that change significantly following CS exposure. Of the 6 significantly regulated loci, 4 map to lysozyme family genes in a region within an intron of the multiple wing hairs (mwh) gene. In some cases these esiRNA producing lysozyme family genes overlap each other on opposite strands. By examining predicted secondary structures of lysozyme mRNAs, and of the mwh intron, I show that esiRNAs produced from these loci are likely derived from long hairpin structures within the mwh intron, and not from overlapping transcription of the lysozyme genes themselves. Reads derived from this region do not map elsewhere in the genome, and are therefore likely to be involved only in regulation of the lysozyme family genes, or of mwh. The mwh intron from which these esiRNAs are produced is not included in all mwh transcripts, suggesting that mwh promoter choice may influence the expression of lysozyme genes.(*32*) Lastly, I identify piRNA producing loci, and examine piRNA expression during LTM formation. I find that no piRNA loci are upregulated, and 10 are downregulated during LTM formation. Invader3 and Dm88 transposons map within downregulated piRNA loci, both of which are long terminal repeat (LTR) type retrotransposons. This work is, to my knowledge, is the first genome wide analysis of sRNA expression in the context of LTM formation in *Drosophila*.

Results

Section I: Analysis of microRNA expression in the Drosophila head during long-term memory formation

microRNA expression in the Drosophila head during LTM formation

Tissue-specific profiles of microRNA expression have been reported, and the expression of certain microRNAs is restricted to a single tissue or sets of cells.(*33-35*) microRNA expression also responds to environmental stimuli and signaling cascades.(*36*) Some microRNAs have been shown to be regulated by neural activity, and the expression of a limited set of microRNA precursors or individual mature microRNAs has been shown to change during LTM formation.(*37-39*) However, genome wide regulation of mature microRNAs has not been studied in the context of *Drosophila* behavior or memory paradigms. To obtain a genome wide profile of microRNA expression in the *Drosophila* brain and to understand changes in microRNA expression during LTM formation, we conducted a high throughput sequencing study of 15-35nt RNAs extracted from heads of flies that had been subjected to aversive olfactory classical conditioning. Training followed the olfactory LTM conditioning paradigm described by Tully and Quinn(*13*) In order to understand the contributions of the US (spaced shocks) and CS (odor only) to changes in microRNA expression following conditioning, we also sequenced 15-35nt RNAs extracted from heads of flies that had been subjected to training identical to

that of LTM conditioned flies, but received either the US alone, the CS alone, or neither (control). We chose to use whole head lysates in this study as opposed to exclusively brain tissue because our primary goal was to identify changes in sRNA expression during LTM formation, rather than simply identifying all microRNAs expressed in the brain alone. The brain and eyes comprise most of the volume of the *Drosophila* head, though muscle, hemolymph, tracheal, and other tissues are also present. We reasoned that the variability introduced by the delicate and time consuming dissection of the brain from the fly head would be more problematic than inclusion of non-neuronal tissues. This also made it possible to greatly increase the number of flies represented in each sequencing library, and as a result, improved the uniformity of RNA extracts used to prepare sequencing libraries. We took great care to ensure that flies in each treatment group differed only in the stimuli to which they were exposed during training. Fly husbandry and handling were carried out meticulously to ensure that all flies used in this work had as closely matching life experiences prior to training as is possible. For each matched set of samples, a single age matched set of flies was split into the four treatment groups. All four treatments occurred simultaneously, and in close proximity, through the use of an automated olfactory conditioning apparatus developed in our lab.(*40*) This apparatus is under computer control, and produces highly replicable training sessions. Heads from 10 male and 10 female flies in each condition were dissected and flash frozen 2 hours after conditioning. Conditioned flies were tested for memory of the pairing of shock and odor 24 hours after conditioning. Sequencing libraries were prepared only from matched treatment sets in which flies in the LTM

condition displayed memory formation 24 hours after conditioning. Because of the great care we took in fly handling, training, and sequencing library preparation, we feel confident in asserting that, despite including tissues of the head other than the brain in our samples, any observed differences in sRNA expression between treatment groups at least primarily reflect differences in neural activity induced by the stimuli used in each treatment group.

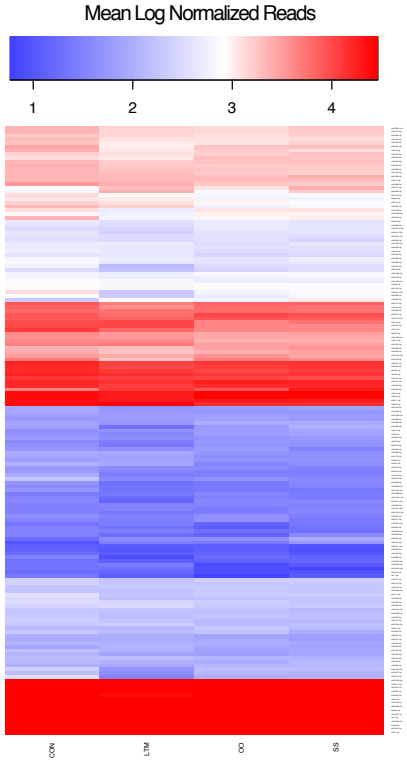
Sequencing was conducted on the Illumina Genome Analyzer IIx platform. Sequencing reads were filtered for quality, and adapter sequences were removed using the FASTX toolkit. Sequencing reads were aligned to the pre-microRNA sequences annotated in miRBase. Counts of reads that aligned to pre-microRNAs were normalized using the upper quartile method of edgeR. Reads that precisely matched the mature microRNA sequences annotated in miRBase were designated as canonical mature microRNAs. While many sequencing reads other than canonical mature microRNA sequences also aligned to pre-microRNAs, we left such reads out of our initial analysis for several reasons. Non-canonical microRNA reads are diverse in origin, span the entire 15-35nt size range we sampled, and are therefore sometimes incompatible with Ago1 loading. Thus, the functional significance of non-canonical microRNA reads may vary greatly, and their inclusion in this part of our analysis would have hindered interpretation. Existing microRNA target predictions are available only for canonical mature microRNAs, thereby restricting downstream analysis of the consequences of any changes in microRNA expression detected. Lastly, previous studies of microRNA function and regulation in the

context of synaptic plasticity and memory have examined only canonical mature microRNAs. Therefore interpreting our results in the context of previous work would be impossible if non-canonical microRNA reads were included. We will discuss our detailed analysis of non-canonical microRNA reads later in this section.

We detected 317 of the 426 mature microRNAs found in the miRBase microRNA database.(Figure S2.1) MiRBase has been curated with data from the modencode consortium, which has used a very large set of publicly available sequencing data to identify and validate candidate microRNAs. Thus, the miRBase catalog of *Drosophila* microRNAs is considered to represent the complete set of microRNAs expressed in flies. We found that a handful of mature microRNAs represent the vast majority of all sequencing reads mapping to microRNA precursors.(Figure 2.1)

Figure 2.1. The 10 most highly expressed mature microRNAs comprise ~90% of all mature microRNA reads

A.



B.

The 10 most abundant mature microRNAs as fractions of all normalized mature microRNA reads

miRNA	Con	LTM	OO	SS	Mean
miR-184-3p	0.3645	0.2932	0.5728	0.5560	0.4466
miR-1-3p	0.3107	0.3651	0.2514	0.2397	0.2917
miR-8-3p	0.0507	0.0523	0.0296	0.0353	0.0420
miR-276a-3p	0.0471	0.0454	0.0172	0.0214	0.0328
bantam-3p	0.0312	0.0346	0.0182	0.0193	0.0258
let-7-5p	0.0263	0.0264	0.0236	0.0249	0.0253
miR-957-3p	0.0261	0.0300	0.0154	0.0176	0.0223
miR-263b-5p	0.0160	0.0199	0.0111	0.0115	0.0146
miR-8-5p	0.0118	0.0146	0.0067	0.0084	0.0104
miR-193-5p	0.0111	0.0129	0.0048	0.0058	0.0087

Sequencing was conducted on 15-35nt RNAs isolated from heads of drosophila in each of the treatment conditions (Con: Control, LTM: Odor + Shock, OO: Odor only, SS: Shock Only). Sequencing reads were aligned to the pre-microRNA sequences annotated in miRBase. Counts of reads that aligned to pre-microRNAs were normalized using the upper quartile method of edgeR. Reads that precisely matched the mature microRNA sequences annotated in miRBase were designated as canonical mature microRNAs. A: heatmap of mature microRNA expression for mature microRNAs with at least 10 reads amongst all samples. B: The top 10 canonical mature microRNAs contributed ~90% of all mature microRNA reads in all conditions.

We used the edgeR differential expression analysis software package to identify mature microRNAs that exhibit changes in expression following treatment. edgeR incorporates functionality that includes batch effects and differing library sizes, in addition to treatment groups in its analysis. This functionality, when combined with our paired sample experimental design, yields a powerful method for detection of truly differentially expressed microRNAs, while minimizing false positives. Recently, our selection of a paired sample experimental design and choice of edgeR was shown to be the best available methodology for differential expression analysis from sequencing data.(*41*) Using this approach, we compared expression of mature microRNAs following LTM conditioning, odor only training (OO), or shock only training (SS), to expression in the control group. Broadly speaking, the profile of LTM condition most strongly differed from controls, while the OO and SS conditions showed fewer changes. Many changes in mature microRNA expression were observed in the various treatment conditions, but only 5 mature microRNAs displayed statistically significant changes in any treatment group.(Figure S2.2) We identified one significantly downregulated mature microRNA, and 4 significantly upregulated mature microRNAs following treatment.(Figure 2.2)

Figure 2.2: Mature microRNAs with statistically significant changes in expression following treatment

	LTM		OO		SS	
microRNA	Log Fold Change	FDR adj. pVal	Log Fold Change	FDR adj. pVal	Log Fold Change	FDR adj. pVal
miR-958-3p	2.314706243	8.01E-05	0.186303652	1	1.265770139	0.357331616
miR-958-5p	2.158785348	0.000363861	1.090557475	1	1.897431458	0.043614936
miR-956-3p	1.812438154	0.008384943	0.394817835	1	1.490435064	0.143442265
miR-314-3p	1.443989577	0.040044656	0.548143742	1	1.477616407	0.185118642
miR-312-3p	-1.659199692	0.042035678	-0.705716964	1	-1.318339372	0.832161977

We identified 5 mature microRNAs with significantly changed expression following treatment. Changes following shock only training resembled those produced by LTM conditioning, while changes produced by odor only training were smaller and not statistically significant. All 5 regulated mature microRNAs responded significantly to LTM conditioning, while only the change in miR-958-5p expression was statistically significant in the shock only condition.

The microRNAs with statistically significant changes in expression following LTM showed similar changes in the SS condition, but only miR-958-5p changed significantly in SS. No mature microRNAs showed statistically significant changes in expression in the OO condition.

Target analysis for microRNAs regulated during LTM formation

Having identified mature microRNAs that are significantly regulated during LTM formation, we examined the downstream consequences of these changes. We used microRNA target predictions from the DIANA microT algorithm, as it uses modern prediction methods, and scans entire mRNAs for target sites as opposed to restricting the search to the 3'UTR as is the case with all other publicly available target prediction databases.(*24*) The full catalog of predicted targets for microRNAs regulated following LTM conditioning is available in Figure S2.3. Together, the

significantly regulated microRNAs are predicted to target transcripts of 1090 genes. Using the microT target predictions, we sought to delineate genes targeted by more than one of the microRNAs regulated following LTM conditioning from those targeted by only one. miR-312-3p has far more predicted targets than any of the other significantly regulated microRNAs. miR-312-3p is predicted to target transcripts of 562 genes, of which 484 are not targeted by any other significantly regulated microRNA. Significantly upregulated microRNAs together target 606 genes, of which 528 are not also targeted by miR-312-3p. 49 genes are targeted by more than one up-regulated microRNA. Of those, 10 genes were also targeted by the lone significantly down-regulated microRNA, miR-312-3p (Figure 2.3).

Figure 2.3. Genes targeted by 2 or more upregulated microRNAs

FlyBase ID	Symbol	Name	dme-miR-312-3p	dme-miR-314-3p	dme-miR-956-3p	dme-miR-958-3p	dme-miR-958-5p
FBgn0003507	srp	serpent	+	+	-	+	-
FBgn0261986	RASSF8	NA	+	+	+	-	-
FBgn0029761	SK	small conductance calcium-activated potassium channel	+	+	+	-	+
FBgn0039078	CG4374	NA	+	+	+	-	-
FBgn0035229	CG7852	NA	+	+	-	+	-
FBgn0030680	CG8944	NA	+	+	+	-	-
FBgn0032341	Reps	NA	+	-	+	+	-
FBgn0262738	norpA	no receptor potential A	+	+	+	-	-
FBgn0038341	CG14869	NA	+	+	+	-	-
FBgn0011582	Dop1R1	Dopamine 1-like receptor 1	+	+	+	+	-
FBgn0024234	gbb	glass bottom boat	-	+	+	-	-
FBgn0262579	Ect4	Ectoderm-expressed 4	-	+	+	-	-
FBgn0040388	boi	brother of ihog	-	+	+	-	-
FBgn0023081	gek	genghis khan	-	+	+	-	-
FBgn0037336	CG2519	NA	-	+	+	+	-
FBgn0263218	NA	NA	-	+	+	-	-
FBgn0260499	qvr	quiver	-	+	+	-	-
FBgn0085400	CG34371	NA	-	+	+	-	-
FBgn0004622	Takr99D	Tachykinin-like receptor at 99D	-	+	+	-	-
FBgn0003071	Pfk	Phosphofructokinase	-	+	-	-	+
FBgn0027512	CG10254	NA	-	+	+	-	-
FBgn0004892	sob	sister of odd and bowl	-	+	-	+	-
FBgn0028550	Atf3	Activating transcription factor 3	-	+	+	-	-
FBgn0016059	Sema-1b	Sema-1b	-	+	+	-	-
FBgn0031632	CG15628	NA	-	+	+	-	-
FBgn0036317	CG10948	NA	-	+	-	+	-
FBgn0028506	CG4455	NA	-	+	+	-	-
FBgn0038829	CG17271	NA	-	+	-	+	-
FBgn0038595	CG7142	NA	-	+	+	-	-
FBgn0053558	mim	missing-in-metastasis	-	+	+	-	-
FBgn0030432	CG4404	NA	-	+	-	-	+
FBgn0024963	GluClalpha	GluClalpha	-	+	-	-	+
FBgn0036522	CG7372	NA	-	+	+	-	-
FBgn0036446	CG9384	NA	-	+	-	+	-
FBgn0035085	CG3770	NA	-	+	-	-	+
FBgn0032378	CycY	Cyclin Y	-	+	+	+	-
FBgn0031637	CG2950	NA	-	+	-	+	-
FBgn0016641	PTP-ER	Protein tyrosine phosphatase-ERK/Enhancer of Ras1	-	-	+	-	+
FBgn0040089	meso18E	meso18E	-	-	+	+	-
FBgn0259938	cwo	clockwork orange	-	-	+	+	-
FBgn0036202	CG6024	NA	-	-	+	+	-
FBgn0038890	CG7956	NA	-	-	+	+	-
FBgn0085446	CG34417	NA	-	-	+	+	-
FBgn0262734	Rbp2	RNA-binding protein 2	-	-	+	-	+
FBgn0000395	cv-2	crossveinless 2	-	-	+	-	+
FBgn0037659	Kdm2	Lysine (K)-specific demethylase 2	-	-	-	+	+
FBgn0030758	CanA-14F	Calcineurin A at 14F	-	-	-	+	+
FBgn0032312	CG14071	NA	-	-	-	+	+
FBgn0036732	Oatp74D	Organic anion transporting polypeptide 74D	-	-	-	+	+

Target predictions were obtained from the DIANA microT algorithm for mature microRNAs displaying statistically significant changes in expression following LTM conditioning. 49 genes are targeted by 2 or more significantly upregulated microRNAs. Of these, 10 are also targeted by the lone significantly downregulated mature microRNA miR-312-3p.

Interestingly, the Dop1R1 dopamine receptor is targeted by all significantly regulated mature microRNAs except miR-958-5p. Expression of this gene in the mushroom body is required for olfactory memory formation.(*42*) The small

conductance calcium-activated potassium channel (SK) gene is targeted by all significantly regulated mature microRNAs except miR-958-3p. This gene is involved in courtship STM and LTM.(*43*) Cyclin Y and the poorly conserved gene CG2519 are targeted by miR-314-3p, miR-956-3p, and miR-958-3p.

Gene ontology analysis of targeted genes

While some target genes such as Cyc-Y, SK, Dop1R1 and CG2519 are obvious candidates for regulated silencing during LTM formation, we hoped to identify other, less immediately obvious candidate genes for further study. We also sought a better understanding of the cellular processes potentially regulated by microRNA mediated silencing during LTM formation. To accomplish both of these goals, we obtained gene ontology annotations for genes targeted by significantly regulated microRNAs from the PANTHER database.(*44*) When the list of all genes targeted by significantly regulated microRNAs is examined for pathways whose component genes are over or underrepresented in the list, only the "Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway" annotation is over or underrepresented. In a list of this size, 3.38 genes in this pathway would be expected to be found, while the actual list of targeted genes includes 12 (FDR adjusted P-val = 0.0367) (Figure S2.4) Interestingly, miR-958-5p is the only significantly regulated microRNA that does not target any of the 12 genes in this pathway(Figure 2.4).

Figure 2.4. Targeting of genes in the Gi alpha and Gs alpha mediated pathway by significantly regulated microRNAs

FlyBase ID	Symbol	miR-312-3p	miR-314-3p	miR-956-3p	miR-958-3p	miR-958-5p
FBgn0003371	sgg	-	+	-	-	-
FBgn0250910	Octbeta3R	+	-	-	-	-
FBgn0024814	Clc	+	-	-	-	-
FBgn0011582	Dop1R1	+	+	+	+	-
FBgn0028433	Ggamma30A	-	-	-	+	-
FBgn0000253	Cam	+	-	+	-	-
FBgn0024941	RSG7	-	-	+	-	-
FBgn0015129		+	-	-	-	-
FBgn0051960	CG31960	-	+	-	-	-
FBgn0031995	CG8475	-	+	-	-	-
FBgn0000037	mAcR	+	-	-	-	-

Pathway annotations for all genes targeted by significantly regulated microRNAs were obtained from the PANTHER database. With 12 genes, the "Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway" was overrepresented in the set of targeted genes (FDR Adj P-val 0.0367). Genes targeted by a given microRNA are marked with a "+".

We repeated the ontology analysis on the set of all genes targeted by significantly regulated microRNAs, but this time examined the classes of proteins encoded by targeted genes, their molecular functions, and the biological processes they participate in. (Figure S.2.5). 17 protein class annotation terms are significantly over or underrepresented (FDR Adj P-val < 0.05). These correspond to 14 categories of molecular functions. 29 biological processes are significantly overrepresented (FDR Adj P-val < 0.05) in the set of all genes targeted by significantly regulated microRNAs (Figure 2.5).

Figure 2.5. Overrepresented biological process annotations for all genes targeted by significantly regulated microRNAs

Biological Process	All Drosophila Genes With Annotation	Targeted Genes	Expected Genes	over/under	FDR Adj. P-value
cell communication	1873	262	150.88	+	3.68E-17
cellular process	2904	357	233.94	+	9.96E-16
signal transduction	1739	242	140.09	+	2.63E-15
developmental process	1193	179	96.1	+	1.53E-13
system development	735	121	59.21	+	3.61E-11
neurological system process	1109	162	89.34	+	4.06E-11
transcription	1049	155	84.5	+	6.08E-11
transcription from RNA polymerase II promoter	1048	154	84.42	+	1.12E-10
system process	1210	171	97.47	+	1.21E-10
intracellular signaling cascade	631	107	50.83	+	1.76E-10
ectoderm development	555	93	44.71	+	1.09E-08
regulation of transcription from RNA polymerase II promoter	781	118	62.91	+	1.51E-08
Unclassified	6079	388	489.7	-	5.99E-08
nervous system development	522	87	42.05	+	6.04E-08
cell surface receptor linked signal transduction	835	121	67.26	+	1.02E-07
muscle contraction	188	44	15.14	+	1.50E-07
mesoderm development	454	76	36.57	+	6.84E-07
muscle organ development	215	46	17.32	+	9.53E-07
cell adhesion	476	78	38.35	+	1.05E-06
cell motion	347	60	27.95	+	1.04E-05
cell-cell signaling	517	78	41.65	+	2.94E-05
apoptosis	442	66	35.61	+	3.70E-04
vesicle-mediated transport	526	73	42.37	+	1.33E-03
transmembrane receptor protein tyrosine kinase signaling pathway	93	22	7.49	+	1.98E-03
protein modification process	765	96	61.63	+	2.93E-03
induction of apoptosis	120	25	9.67	+	4.25E-03
nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	2117	218	170.54	+	1.32E-02
MAPKKK cascade	129	25	10.39	+	1.32E-02
synaptic transmission	393	55	31.66	+	1.40E-02
sensory perception of sound	44	13	3.54	+	1.42E-02
embryonic development	168	29	13.53	+	2.67E-02

Biological process annotations for all genes targeted by significantly regulated microRNAs were obtained from the PANTHER database. 29 biological processes are significantly overrepresented (FDR Adj P-val < 0.05) in this set, including the terms "neurological systems process" and "synaptic transmission".

Of note, the terms "neurological systems process" and "synaptic transmission" are significantly overrepresented. All genes annotated with "synaptic transmission" are also included in the "neurological systems process" annotation, as the latter is one of the parent GO terms for genes expressed it neurons or glia. The 154 target genes with this annotation are attractive candidates for subsequent investigation. We therefore further parsed the 154 neurological systems process genes by the microRNAs that target each gene (Figure S2.6). 68 of 154 genes with the "neurological systems process" annotation were uniquely targeted by miR-312-3p, and 68 genes with this annotation were targeted by one or more significantly upregulated microRNA, but not by miR-312-3p. 26 "neurological systems process" genes were targeted by 2 or more significantly regulated microRNAs (Figure 2.6).

Figure 2.6. Genes annotated with the term "neurological systems process" and targeted by multiple significantly regulated microRNAs

FlyBase ID	Symbol	miR-312-3p	miR-314-3p	miR-956-3p	miR-958-3p	miR-958-5p
FBgn0011582	Dop1R1	+	+	+	+	-
FBgn0032341	Reps	+	-	+	+	-
FBgn0000448	Hr46	+	+	-	-	-
FBgn0003870	ttk	+	-	+	-	-
FBgn0004242	Syt1	+	+	-	-	-
FBgn0004622	Takr99D	-	+	+	-	-
FBgn0013733	shot	+	-	-	+	-
FBgn0014870	Psi	+	-	-	-	+
FBgn0015774	NetB	+	-	-	+	-
FBgn0016059	Sema-1b	-	+	+	-	-
FBgn0017549	Ric	+	+	-	-	-
FBgn0024963	GluClalpha	-	+	-	-	+
FBgn0026086	Adar	+	-	+	-	-
FBgn0028550	Atf3	-	+	+	-	-
FBgn0028704	Nckx30C	+	+	-	-	-
FBgn0029508	Tsp42Ea	+	-	-	+	-
FBgn0031760	Tsp26A	+	-	-	+	-
FBgn0034433	endoB	+	-	-	+	-
FBgn0038890	CG7956	-	-	+	+	-
FBgn0039431	CG6490	+	-	+	-	-
FBgn0040388	boi	-	+	+	-	-
FBgn0085446	CG34417	-	-	+	+	-
FBgn0259231	CCKLR-17D1	+	+	-	-	-
FBgn0262350	bru-3	+	+	-	-	-
FBgn0262737	mub	+	-	+	-	-
FBgn0263218	Dscam2	-	+	+	-	-

154 genes targeted by significantly regulated microRNAs are annotated with the term "neurological systems process". Of these, 26 are targeted by more than one significantly regulated microRNA. Genes targeted by a given microRNA are marked with a "+".

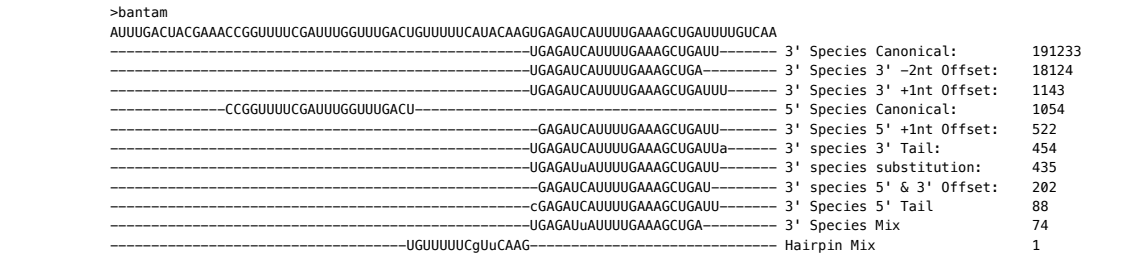
We conducted similar gene ontology enrichment analyses on the sets of predicted targets of each microRNA individually, and for the set of targets unique to miR-312-3p, the set of targets of only upregulated microRNAs, and for the set of targets shared by miR-312-3p and at least one upregulated microRNA. However, these analyses did not yield additional insights as the smaller sets of genes reduced the statistical power of the enrichment analysis, and those annotations that were found to be significantly over or underrepresented were also found in the analysis of the set of all genes targeted by significantly regulated microRNAs.

Expression of non-canonical microRNA sequences in the Drosophila head during LTM formation

In addition to canonical mature microRNAs, a variety of other sequencing reads map to pre-microRNAs. Previous studies have suggested that such non-canonical reads may be indicative of altered microRNA processing, stability, or RNA editing activity. Many non-canonical reads, which are known as isomiRs, will have substantially different regulatory function than their corresponding canonical mature microRNAs. We therefore sought to profile non-canonical microRNA reads in *Drosophila* heads, and to determine if this profile changes during LTM formation. In order to detect non-canonical microRNA reads, we aligned our sequencing reads to pre-microRNA hairpin sequences from miRBase, such that up to two nucleotide mismatches were permitted. Only alignments with scores equal to the highest alignment score for a given read were retained. This approach permits alignment of reads featuring untemplated or edited nucleotides, while minimizing the chances of spurious alignments being retained. Instances of more than two untemplated nucleotide additions to microRNA reads have been documented by other groups, but such reads would be rejected by our methodology. The computational requirements of such an analysis are beyond the scope of this study, and previous work has demonstrated that reads with one or two untemplated nucleotides dominate tailed species, and that addition of up to two untemplated nucleotides indicates the activation of tailing mechanisms.(*45*) We therefore felt that our approach would capture a representative fraction of tailing events, while not imposing excessive computational burdens. In addition to canonical mature microRNAs, we observed several classes of non-canonical microRNA reads using

our approach. 3' offset reads have a canonical 5' end, and no untemplated nucleotides, but are either shorter or longer than the canonical mature microRNA. Similarly, 5' offset reads have canonical 3' ends and no untemplated nucleotides, but differ from canonical mature microRNAs at their 5' end. Overlap reads have no untemplated nucleotides, and overlap the position of their corresponding canonical mature microRNA, but have noncanonical 3' and 5' ends. 3' tailed reads have canonical 5' ends, may have a noncanonical 3' end, and up to 2 untemplated 3' nucleotides. 5' tailed reads have canonical 3' ends, and up to 2 untemplated nucleotides at the 5' end. Substitution reads contain up to 2 untemplated nucleotides at positions more than 2 nucleotides away from both the 5' and 3' ends. Hairpin loop reads map to the region between the 5' and 3' canonical species derived from a given pre-microRNA. Mix reads exhibit more than one of the above feature types(Figure 2.7).

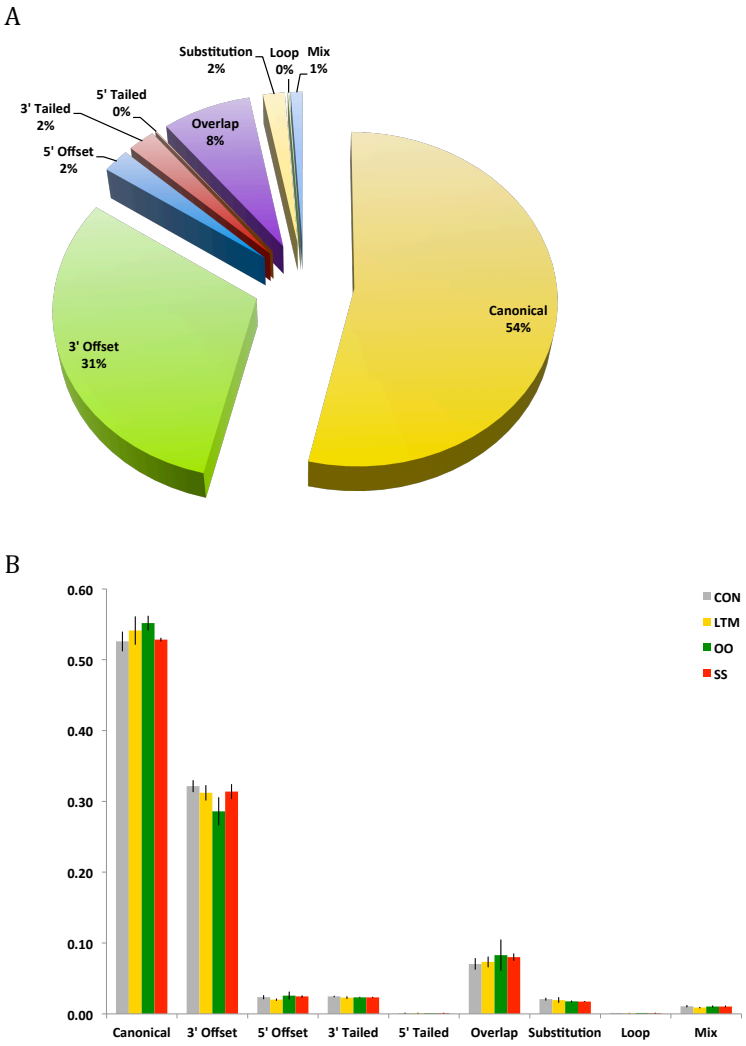
Figure 2.7. Examples of isomiR classes.



Sequencing reads were aligned to pre-microRNA hairpins such that up to 2 mismatches were permitted. Only alignments with alignment scores equal to the best score for a given read were retained. Read counts were upper quartile normalized. 3' offset reads have a canonical 5' end, and no untemplated nucleotides, but are either shorter or longer than the canonical mature microRNA. 5' offset reads have canonical 3' ends and no untemplated nucleotides, but differ from canonical mature microRNAs at their 5' end. Overlap reads have no untemplated nucleotides, and overlap the position of their corresponding canonical mature microRNA, but have noncanonical 3' and 5' ends. 3' tailed reads have canonical 5' ends, may have a noncanonical 3' end, and up to 2 untemplated 3' nucleotides. 5' tailed reads have canonical 3' ends, and up to 2 untemplated nucleotides at the 5' end. Substitution reads contain up to 2 untemplated nucleotides at positions more than 2 nucleotides away from both the 5' and 3' ends. Hairpin loop reads map to the region between the 5' and 3' canonical species derived from a given pre-microRNA. Mix reads exhibit more than one of the above feature types.

Canonical mature microRNAs dominate microRNA reads. However, 3' offset reads represent a significant fraction of reads mapping to pre-microRNAs. None of the treatment conditions we subjected flies to produced a significant change in the proportions of all pre-microRNA mapping reads that fell into each category (Figure 2.8).

Figure 2.8. Noncanonical microRNA reads as a proportion of all reads mapping to pre-microRNAs in each condition



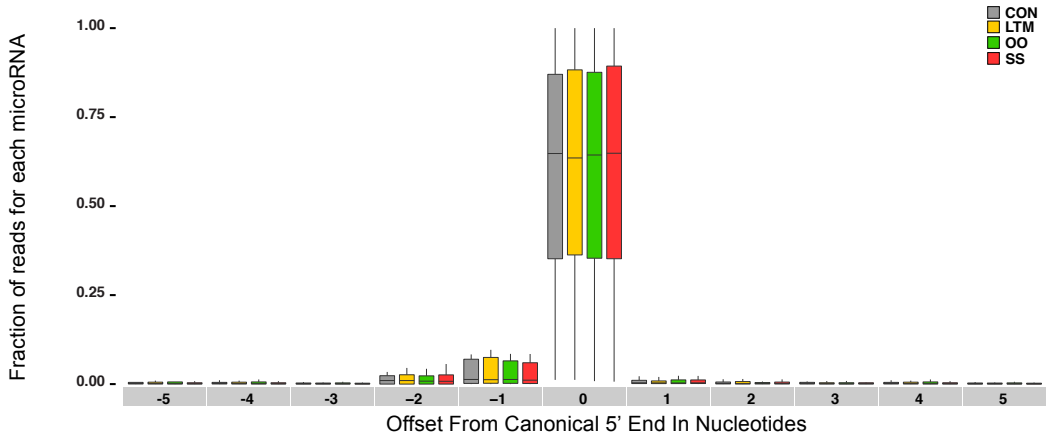
The fraction of each category of noncanonical microRNA reads was calculated for each library. A. Mean fractions of normalized reads from all conditions falling into each isomiR category. B. proportions of each isomiR class for each treatment condition. No treatment condition produced a statistically significant change in the proportion of all pre-microRNA mapping reads falling into any of these categories. Error bars represent SEM.

Offset isomiR analysis for individual pre-microRNAs

Having found no significant changes in the overall proportions of the various classes of isomiR reads amongst all pre-microRNA mapping reads, we next sought to

investigate isomiR read features in detail. As microRNA target specificity is largely determined by the 5' seed sequence, any changes in the position of the 5' end of a microRNA will greatly affect the set of mRNAs it targets. In general, microRNA processing is highly precise at the 5' end. Of those microRNAs with a substantial fraction of 5' isomiR reads, most 5' offset isomiRs are offset by 1-2nt (Figure 2.9).

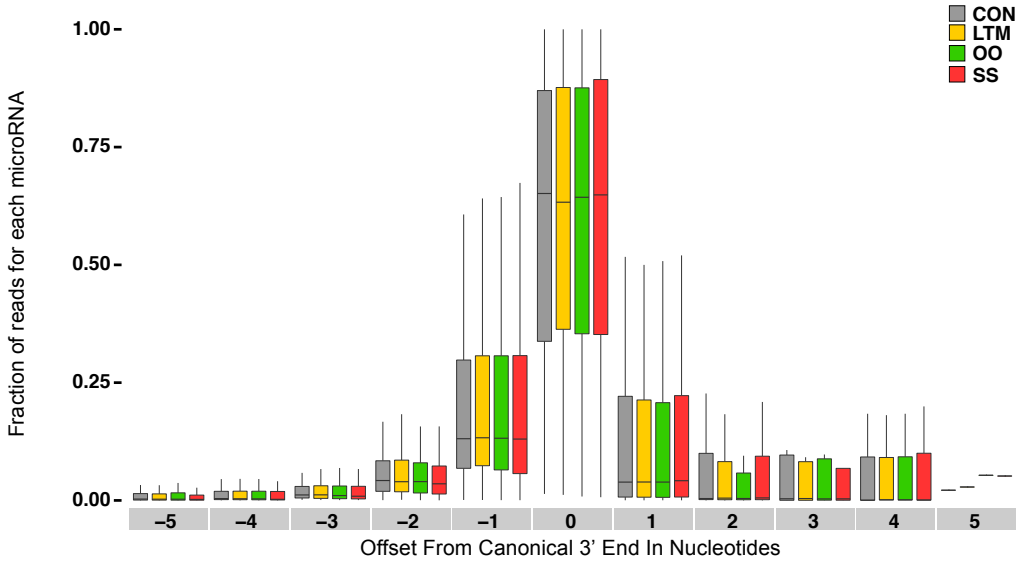
Figure 2.9. Proportion of 5' Offset isomiR reads for each mature microRNA



For each mature microRNA, the proportion of reads with a 5' end at each nucleotide position was calculated. The overwhelming majority of microRNA reads have a canonical 5' end. Of those with a 5' offset, reads extending 1 or 2 nucleotides 5' of the canonical 5' position are most common. Rectangles mark the 25th and 75th percentile. Whiskers are 5th and 95th percentile, dots denote outliers.

A comparison of 5' isomiR fractions for each microRNA in all of our treatment groups shows that no treatment produces a statistically significant change in the fraction of 5' isomiR reads for any microRNA. While the 5' ends of microRNA reads overwhelmingly correspond to the canonical 5' position, substantial diversity in the 3' ends of microRNA reads exists. 25% or more of microRNA reads have noncanonical 3' ends for 105 mature microRNAs. For most microRNAs, 3' isomiRs are primarily truncated by 1nt. However, extension of the 3' end by as much as 4nt are not uncommon for many microRNAs (Figure 2.10).

Figure 2.10. Proportion of 3' Offset isomiR reads for each mature microRNA



For each mature microRNA, the proportion of reads with a 3' end at each nucleotide position relative to the canonical position was calculated. While most microRNA reads have a canonical 3' end, a substantial fraction of reads do not Of those with a 3' offset, reads extending 1 nucleotide 5' or 1 nucleotide 3' of the canonical 3' position are most common. Rectangles mark the 25th and 75th percentile. Whiskers are 5th and 95th percentile, dots denote outliers.

Analysis of untemplated nucleotide tailing

Previous studies have demonstrated that untemplated nucleotide addition affects the stability and function of microRNAs. While the 5' seed sequence largely determines target specificity for microRNAs, 3' nucleotide addition is thought to stabilize or mark microRNAs for destruction. Specifically, 3' adenylation is thought to stabilize animal microRNAs and affect their targeting, while 3' uridination is thought to destabilize microRNAs (*35*, *46-49*) We therefore sought to profile such modifications to microRNA reads in our libraries. Using the same alignment method as was used for the offset analysis, we calculated the fraction of all reads mapping to each mature microRNA locus that feature 3' untemplated adenines or uridines. For

our analysis of 3' untemplated nucleotide addition, we considered only microRNAs that had more than 100 normalized reads per sample. We found that 16 microRNAs had 1% or more of their reads 3'adenylated in any condition, and only miR-927-3p has more than 20% of its reads 3' adenylated in any condition (Figure 2.11 A).

Figure 2.11. 3' Tailing of microRNA reads

A.

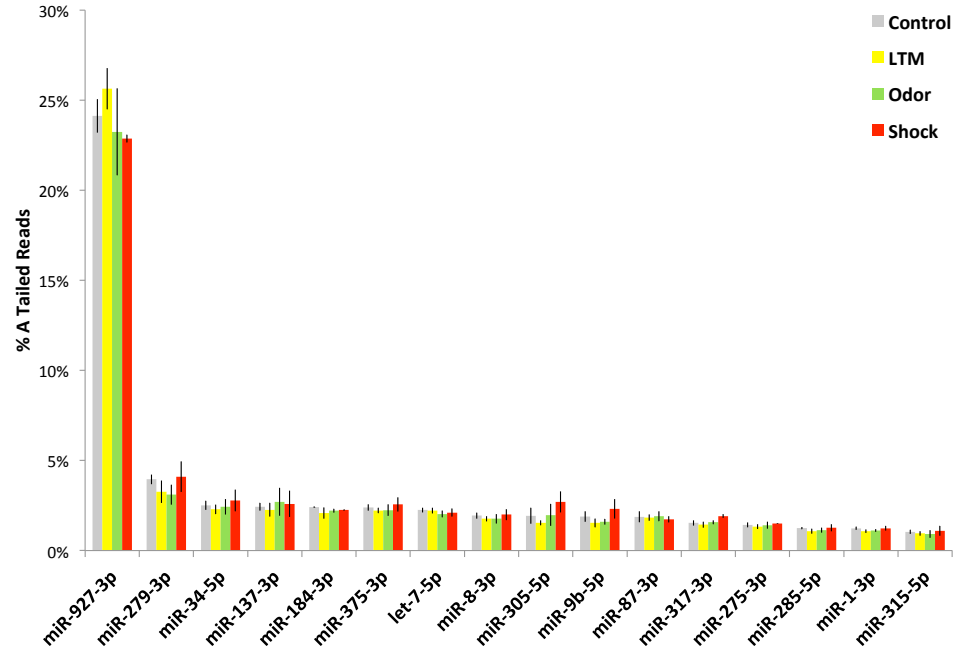
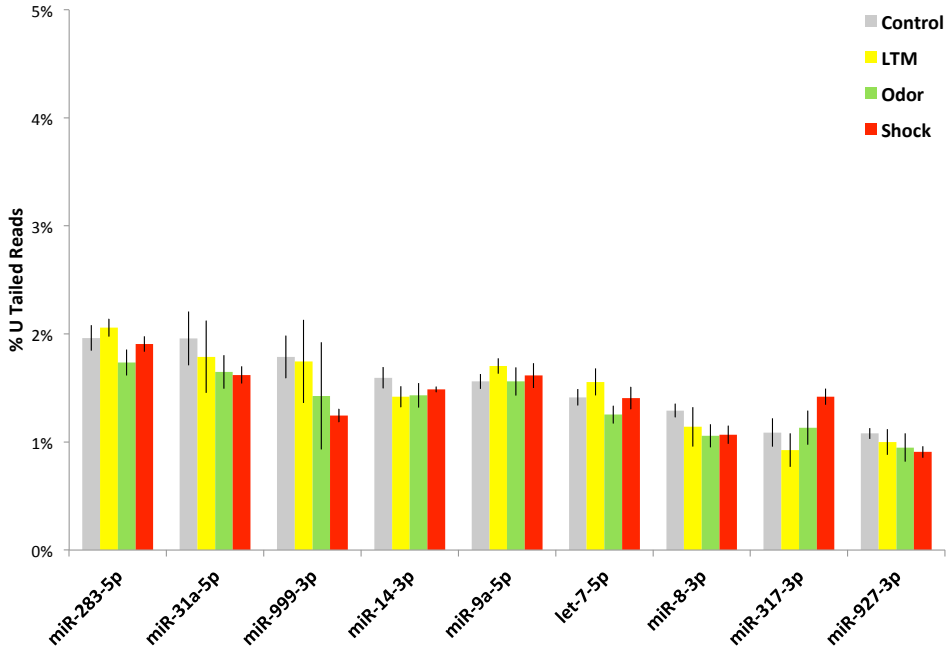


Figure 2.11 (Continued)
B.



Reads mapping at mature microRNA loci were examined for 3' untemplated nucleotide addition. Read counts were upper quartile normalized, and only microRNAs with an average of 100 or more normalized reads per sample were considered in our analysis of 3' tailing. A. 16 microRNAs had greater than 1% of their reads featuring mono or polyadenylation in any treatment condition. B. 9 microRNAs had greater than 1% of their reads featuring mono or polyuridination in any treatment condition.

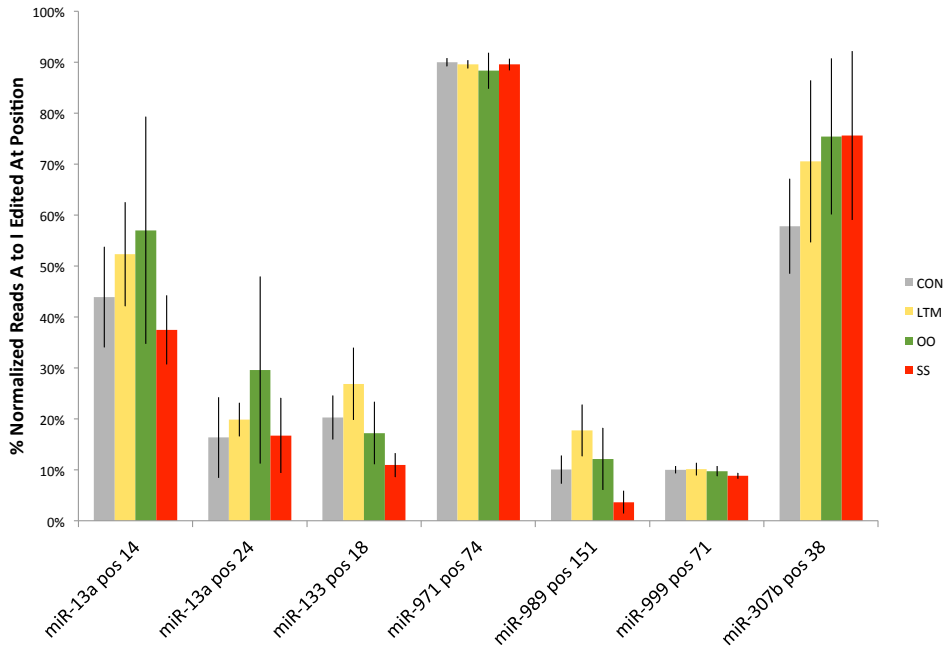
We found that 4 microRNAs had 2% or more of their reads 3' uridinated in any condition, and only miR-970-3p has more than 10% of its reads 3' uridinated in any condition (Figure 2.12 B). Interestingly, miR-927-3p also had greater than 1% of its reads 3' uridinated. We also looked for statistically significant changes in the fractions of 3' adenylated or 3' uridinated reads for all microRNAs. No microRNA displayed a statistically significant change in 3' A or U tailing in any of our treatment conditions. A similar examination of 5' untemplated additions to microRNA reads found such modifications to be essentially absent in our libraries.

Analysis of microRNA editing

Previous studies have found that the RNA editing enzyme adenosine deaminase acting on RNA (ADAR) is highly expressed in the *Drosophila* CNS, and particularly in the MB. Further, ADAR interacts with RISC components, and regulates synaptic structure in *Drosophila*. We therefore examined our sequencing libraries for evidence of pre-microRNA or microRNA editing. Again, we aligned our sequencing reads to pre-microRNA hairpin sequences from miRBase, such that up to two mismatches were allowed, and only alignments with scores equal to the best alignment score for a given read were retained. We then examined instances in which untemplated nucleotides occurred greater than 2nt away from both the 3' and 5' ends of a read. This step was to ensure that we did not include tailed microRNA reads in our analysis, as terminal untemplated nucleotide addition occurs via a mechanism distinct from ADAR activity. Illumina sequencing technology reads inosine nucleotides as guanine. If nearly all of our reads feature a nucleotide that differs from the reference sequence at a given position, we would consider this to be a single nucleotide polymorphism, and not the result of ADAR activity. Conversely, if few reads feature a nucleotide that differs from the reference sequence at a given position within a pre-microRNA, we would consider this to be the result of sequencing noise. Thus, we considered instances in which between 10% and 90% of reads mapping to a given pre-microRNA nucleotide featured an A to G substitution to be a genuine editing event. Further, we required at least 15 normalized reads to have mapped to the position in all libraries in order to include

the position in our editing analysis. Using these criteria, we identified 7 genuine editing events within pre-microRNA sequences (Figure 2.12).

Figure 2.12. A to I editing instances within pre-microRNAs



Sequencing reads were mapped to pre-microRNA sequences from miRBase such that up to 2 mismatches were permitted. Only mismatches at least 2nt from both the 5' and 3' ends of the read were considered. The fraction of reads mapping at each position along pre-microRNAs that featured an A to G substitution at that position was computed. Positions with at least 15 normalized reads mapped in all libraries and having between 10% and 90% of reads featuring an A to G substitution at that position were considered genuine editing events. Error bars represent SEM.

Having identified several apparently genuine editing events, we next examined our data for evidence that any of our treatments caused a change in A to I editing rates at these positions. Using a student's t-test, we found that none of the observed differences between conditions in the fraction of A to I edited reads at a given position were statistically significant.

Section II: Analysis of esiRNA and piRNA expression in the Drosophila head during long-term memory formation

The importance of microRNA mediated translational control in synaptic plasticity and memory formation is well established. Recently, several publications have indicated that esiRNAs and piRNAs are involved in these processes as well in nematodes, mollusks, and mice. (*5*, *26*, *50-52*) piRNAs and esiRNAs are expressed in the *Drosophila* head, leaving open the possibility that these classes of sRNAs could be involved in synaptic plasticity and memory formation in flies as well.(*53*, *54*) Though several studies have demonstrated their presence in *Drosophila* heads, none have directly addressed whether esiRNAs or piRNAs are involved in memory formation. More broadly, the potential roles of esiRNAs and piRNAs in memory formation in any animal species remain poorly understood. We therefore sought to profile expression of these sRNA classes in *Drosophila* heads, and to identify any changes in these profiles during memory formation. With these goals in mind, we constructed the sRNA sequencing libraries described in section I of this chapter such that we could capture esiRNAs and piRNAs in addition to microRNAs. The catalog of microRNAs expressed by *Drosophila* is well understood and thoroughly curated.(*35*, *54*) However this is not the case for other sRNA classes. Furthermore, while microRNAs are highly conserved, and transcribed from known genomic locations, esiRNAs and piRNAs are generated from less well defined regions within repetitive elements, mobile elements, and in concert with heterochromatin

formation. Thus, studying these classes of sRNAs demands computational approaches that differ from those for microRNAs.

Identification of esiRNA producing loci

To study esiRNA and piRNA expression, we first had to identify loci that produce these sRNAs. To eliminate loci that could potentially confound our analysis, we first filtered our sequencing data for reads mapping perfectly to pre-microRNA hairpins, tRNAs, rRNAs, snRNAs, and snoRNAs. We then mapped the filtered reads to the *Drosophila* reference genome available from FlyBase such that no mismatches were allowed, and all perfect alignments to the genome were retained.(*32*) To capture all possible loci, we merged the combined alignments from all of our sequencing libraries into 'read-contigs'. All alignments that overlap or are immediately adjacent to each other form a read-contig. This process yielded 874802 read-contigs. However, this number is inflated, as many identical loci are produced by repetitive elements. To reduce this number, and to generate more useful read-contigs, we further merged all read-contigs within 100nt of each other, yielding 219482 read-contigs. This number still includes read-contigs comprising a single read in one library. To further refine our search, we filtered our read-contigs such that only read-contigs with at least 250 reads in our combined sequencing libraries were retained. This criteria translates to roughly 15 reads per locus per library. As reads within 100nt of each other are bridged, this threshold is still quite low. We used this set of filtered read-contigs as the basis for subsequent analysis. In addition to other

characteristics, esiRNAs and piRNAs are known to have distinctive size ranges. esiRNAs are almost entirely 21-22nt in length. Therefore, we identified read-contigs in which at least 75% of reads from all libraries are 21-22nt as esiRNA loci. This approach yielded 368 esiRNA loci (Figure 2.13).

Figure 2.13. Properties of read-contigs

A

Contig Type	Count	Avg. Contig Length	SEM Contig Length
All Contigs	219482	190.74	1.36
All Contigs >= 250 Reads	7581	2176.43	28.83
esiRNA (21-22nt)	368	4536.40	231.73

B

	Total	Intron	Exon	5' UTR	3' UTR	Intergenic	Transposon	Gene	Repeat	Gene w/o Transposon	Gene w/o Repeat
# of esiRNA Loci	368	192	51	71	41	177	214	212	354	67	14
Fraction of all esiRNA Loci	100.00%	52.17%	13.86%	19.29%	11.14%	48.10%	58.15%	57.61%	96.20%	18.21%	3.80%

Read-contigs were generated by merging all sequencing reads from all of our libraries whose ends fall within 100nt of each other into a single contig. We refined this list of read-contigs by retaining only those with at least 250 reads from our combined libraries. esiRNA loci were defined as those read-contigs with >= 75% of mapped reads from our combined libraries being 21-22nt in length. A: esiRNA loci are longer than read-contigs. B: The types of genomic features overlapping each esiRNA locus were obtained using bedtools to intersect feature position lists with esiRNA loci positions. We then calculated the fraction of esiRNA loci overlapping each type of feature.

esiRNA expression profile in the Drosophila head

Closer inspection of the esiRNA loci we identified reveals a variety of sources for esiRNAs, each with unique features (Figure 2.14).

Figure 2.14 examples of esiRNA loci

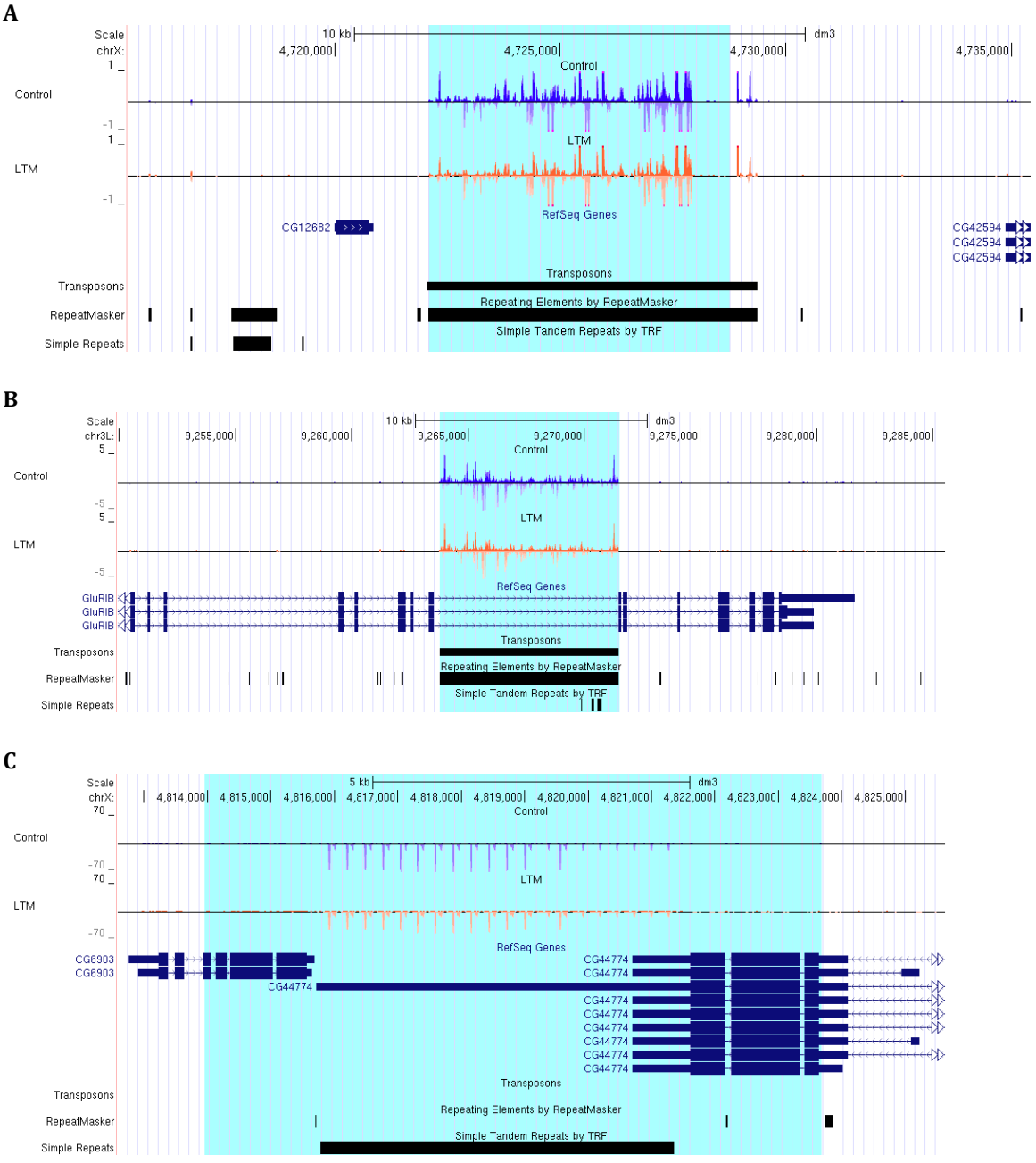
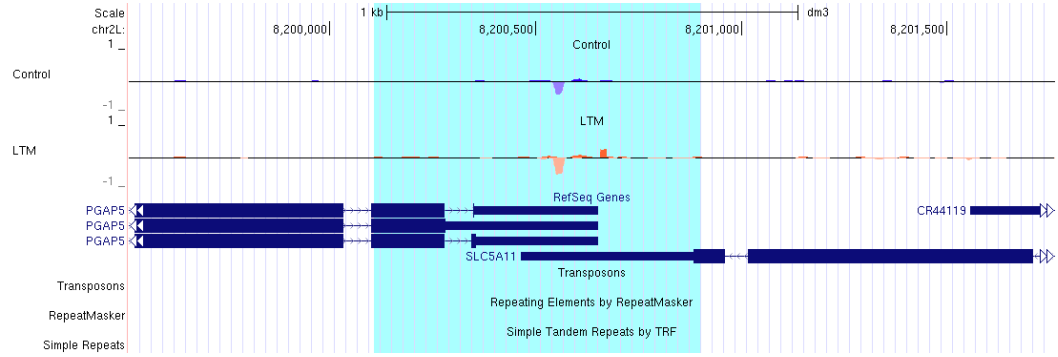


Figure 2.14 (Continued)
D



Reads were mapped to the Drosophila genome such that no mismatches were permitted, and all perfect alignments were retained. Read counts were upper-quartile normalized across all libraries, and average normalized read counts for each condition were computed at every nucleotide position of the genome. These average read depth values were then plotted across the genome using the UCSC genome browser. Read depth values above the black line in each condition represent alignments on the + strand, while values below the line represent alignments on the – strand. Numeric read depth values on the Y-axis represent per-condition mean normalized reads. The light blue shaded region indicates the location of the esiRNA locus. A: Example of an esiRNA locus that maps to an isolated transposon. B: Example of an esiRNA locus that maps to a transposon within an intron of the GluR1B gene. C: Example of an esiRNA locus that maps to a long tandem repeat in the 3' UTR of the CG44774 gene and the adjacent CG6903 gene. These genes are transcribed from opposing strands, but do not overlap. D: Example of an esiRNA locus that maps to the Post-GPI attachment to proteins 5 ortholog (PGAP5) and Sodium/solute co-transporter-like 5A11 (SLC5A11) genes. A stretch of overlapping convergent transcription within the 3' UTRs of both genes produces substantially more reads than the surrounding region, and unlike the surrounding region, a substantial number of reads map to both strands.

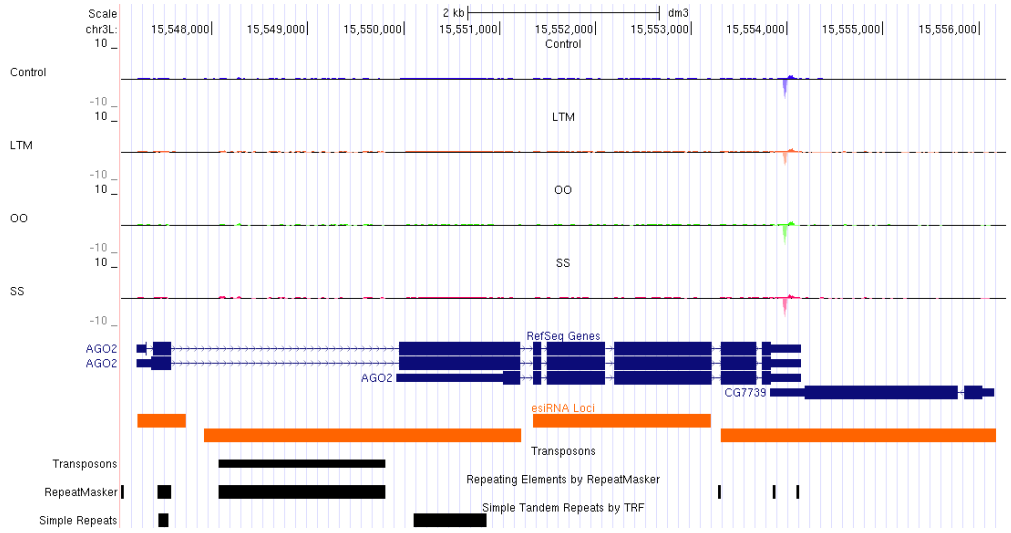
The bulk of esiRNA loci include repetitive elements and transposons, both of which can reside in intergenic regions, or within genes (Figure 2.14 A, B, & C). Such loci produce large numbers of reads, and can span several Kb. In the case of inverted repeats. The esiRNA loci we identified also include regions of overlapping transcription on opposite strands, so called cis-natural antisense transcripts (cis-NATs) (Figure 2.14 D). As reported in other studies, the only examples of cis-NATs that exhibited a strong 21-22nt bias in our libraries are cases in which transcription on the two strand is convergent.(*55*) Our esiRNA loci that map to cis-NATs often extend beyond the region of overlap between the two genes. However, in such cases most of the reads within the esiRNA locus still map within the region of overlap.

The extension of cis-NAT esiRNA loci is due to the presence of reads nearby, but outside of the overlap, and bridging of even single reads within 100nt of each other during the generation of our read-contigs.

Previously reported profiles of esiRNA expression in various *Drosophila* tissues and cell lines have noted several interesting regions that produce esiRNAs (*33*, *54*, *56*-*59*). The esiRNA loci we identified also include many of these same regions. For instance, the production of esiRNAs from the cis-NAT region encompassing the Ago2 and CG7739 genes produces abundant 21-22nt reads. Our methods identify 4 esiRNA loci in this region. Our analysis also identified 2 esiRNA loci mapping to the pseudogene CG18854, which overlaps the Inositol 1,4,5-triphosphate kinase 1 (IP3K1) gene on the opposite strand. CG18854 contains an inverted repeat that produces esiRNAs that were shown to be functional by Okamura et al (Fig 2.15).(*59*)

Figure 2.15. Comparison of previously published esiRNA loci and those identified in this study

A



B

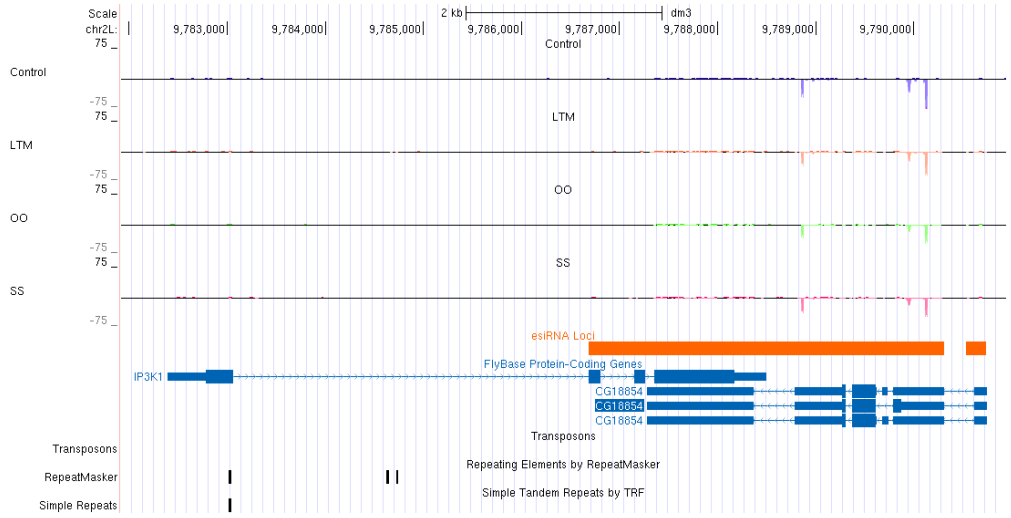
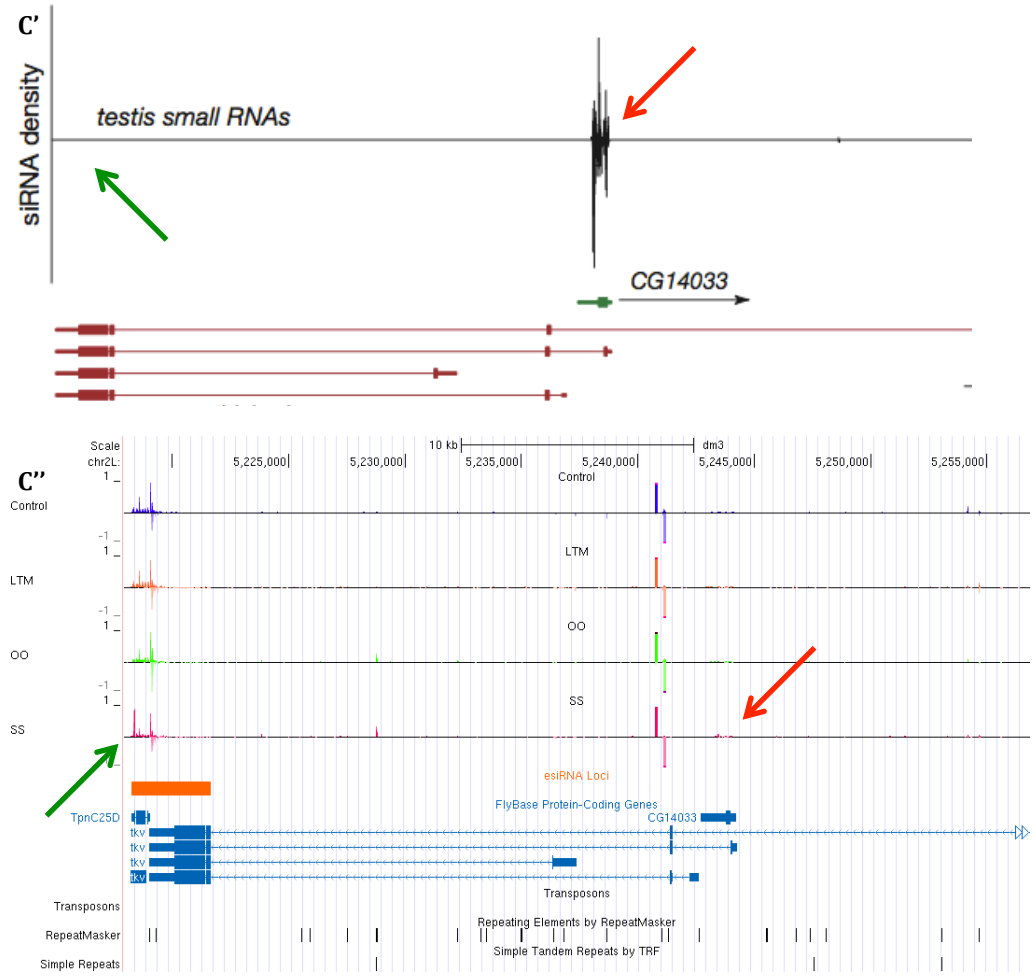


Figure 2.15. (Continued)

C



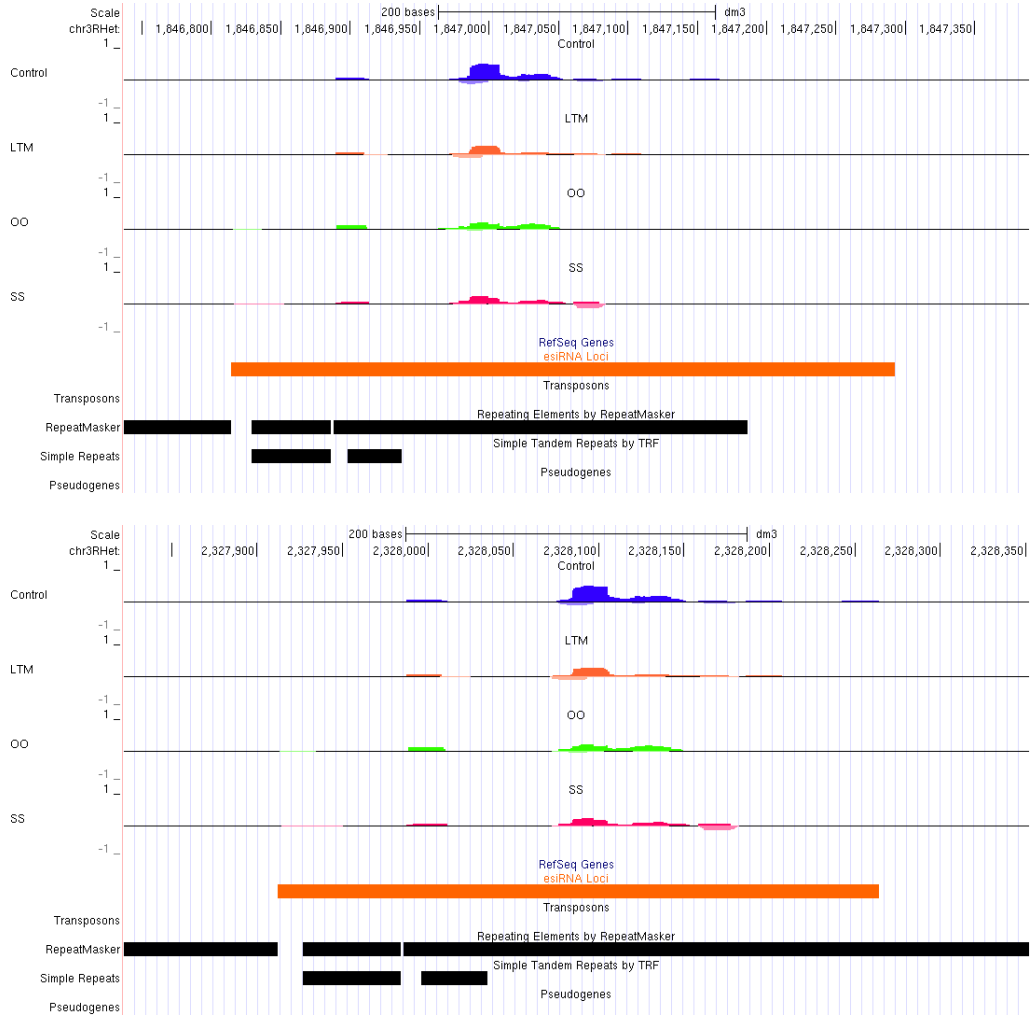
Sequencing reads were aligned to the *Drosophila* genome such that no mismatches were permitted, and all perfect alignments were retained. Read counts were upper-quartile normalized across all libraries, and average normalized read counts for each condition were computed at every nucleotide position of the genome. These average read depth values were then plotted across the genome using the UCSC genome browser. Read depth values above the black line in each condition represent alignments on the + strand, while values below the line represent alignments on the – strand. Numeric read depth values on the Y-axis represent per-condition mean normalized reads. The positions of the esiRNA loci we identified are represented in orange. A: As reported elsewhere, our data shows that region encompassing the esiRNA effector Argonaute protein Ago2 produces abundant esiRNAs from the coding strand. The region of overlap between the Ago2 3'UTR and the CG7739 produces abundant esiRNAs from both strands, a hallmark of esiRNA production from cis-NATs. B: The pseudogene CG18854 overlaps the 3' end of the IP3K1 gene and produces abundant esiRNAs from 2 esiRNA loci. esiRNAs produced from this region are known to be functional. C: The pattern of esiRNA expression at the 3' end of the tkv gene in our libraries differs from that reported elsewhere. C': Czech et al report abundant esiRNA production corresponding to the CG14033 locus (red arrow) and little or no esiRNA production from the 3' tkv exon, which is proximate to TpnC25D (green arrow). C'': In our libraries, few esiRNAs are produced from the CG14033 locus (red arrow), but abundant esiRNAs are produced from the 3' exon of the tkv gene (green arrow).

In other cases, our data shows features not reported elsewhere. For instance, Czech et al. reported abundant esiRNA production from the pseudogene CG14033, which is proximate to the 5' exon of the thickveins-C (tkv-C) transcript in libraries prepared from *Drosophila* testis.(*57*) However, these authors did not report substantial esiRNA production from the 3' end of the tkv gene, which is proximate to the Troponin C at 25D (TpnC25D) gene on the opposite strand. This esiRNA expression pattern is also reported in many of the modencode sRNA tracks. While we see reads at both positions, our data shows that the 3' tkv/TpnC25D locus produces far more reads than does the tkv/CG14033 locus. In fact, the tkv/CG14033 locus produces so few reads, that our methods do not identify it as an esiRNA locus.

Changes in esiRNA expression during LTM formation

We next sought to identify changes in esiRNA expression during LTM formation at the loci we identified. Read counts at esiRNA loci were obtained using BEDTools. We used edgeR in its GLM mode to test esiRNA loci for differential expression. 6 esiRNA loci display statistically significant changes in esiRNA expression during LTM formation. 2 Loci mapping to identical repetitive elements in the heterochromatin of chromosomal arm 3R display significantly reduced read counts in the LTM condition, but not in the odor or shock conditions (Fig 2.16).

Figure 2.16. esiRNA loci downregulated during LTM formation



Sequencing reads were aligned to the *Drosophila* genome such that no mismatches were permitted, and all perfect alignments were retained. Read counts were upper-quartile normalized across all libraries, and average normalized read counts for each condition were computed at every nucleotide position of the genome. These average read depth values were then plotted across the genome using the UCSC genome browser. Read depth values above the black line in each condition represent alignments on the + strand, while values below the line represent alignments on the – strand. Numeric read depth values on the Y-axis represent per-condition mean normalized reads. The positions of the esiRNA loci we identified are represented in orange. Each esiRNA locus was tested for differential expression using edgeR. 2 esiRNA loci mapping to identical repetitive elements in the heterochromatin of chromosomal arm 3R showed significantly reduced read counts in the LTM condition. Although read counts within these loci were also reduced in the odor only (OO) and shock only (SS) conditions, edgeR did not find these changes to be statistically significant.

As these loci are identical in sequence, reads mapping within them map to both loci.

However, these reads do not map elsewhere in the genome. Both down-regulated

loci are separated from known coding and regulatory sequences by tens of Kb. Therefore, these loci and the reads mapping within them do not have obvious regulatory functions. The 4 esiRNA loci displaying statistically significant increased read counts all map to lysozyme family genes residing within a single intron of the multiple wing hairs (mwh) gene (Fig 2.17A).

Figure 2.17. esiRNA Loci mapping to lysozyme family genes within an intron of the mwh gene are significantly up-regulated during LTM formation.

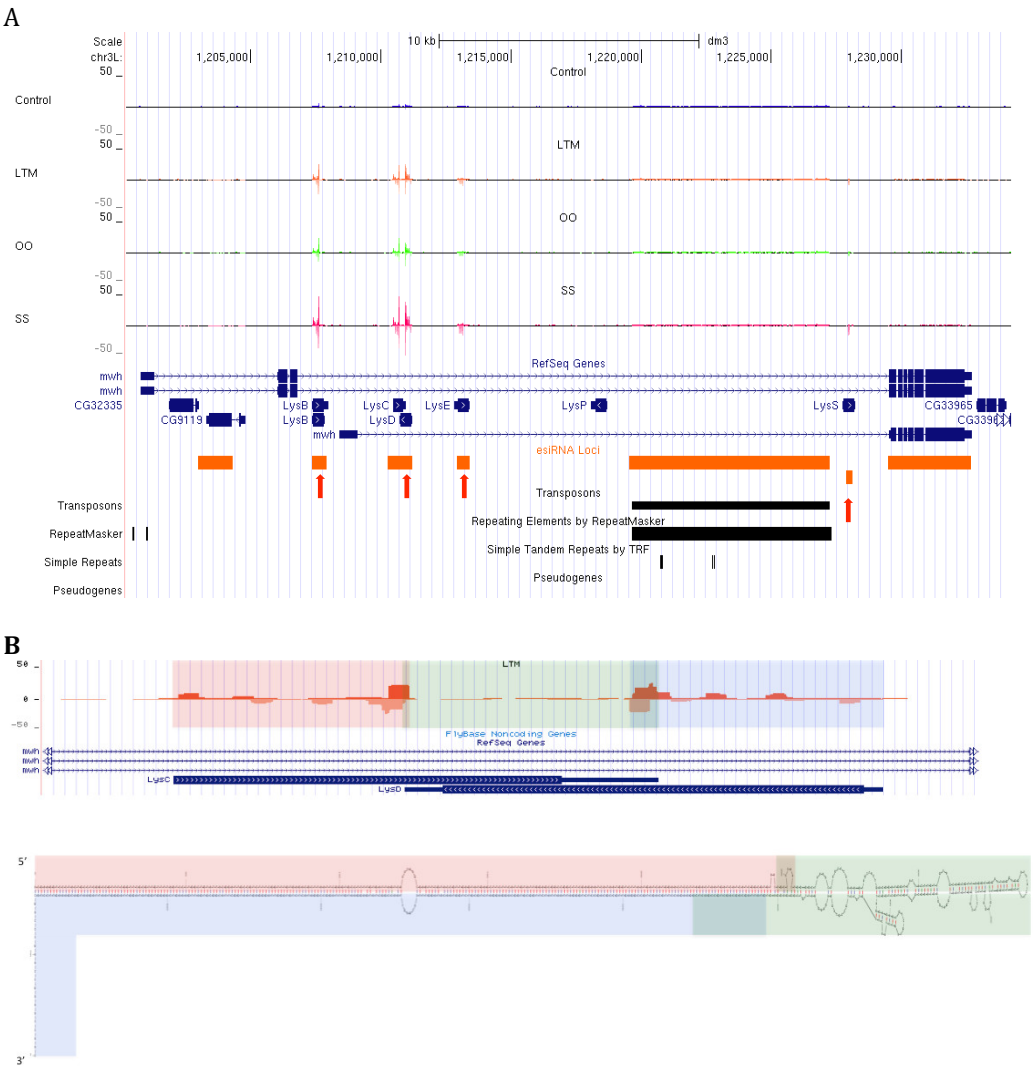
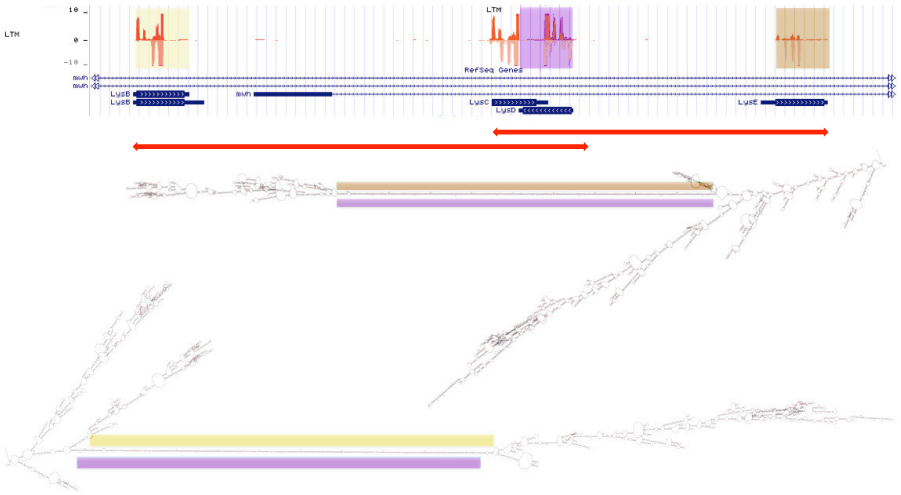


Figure 2.17. (Continued)

C



D

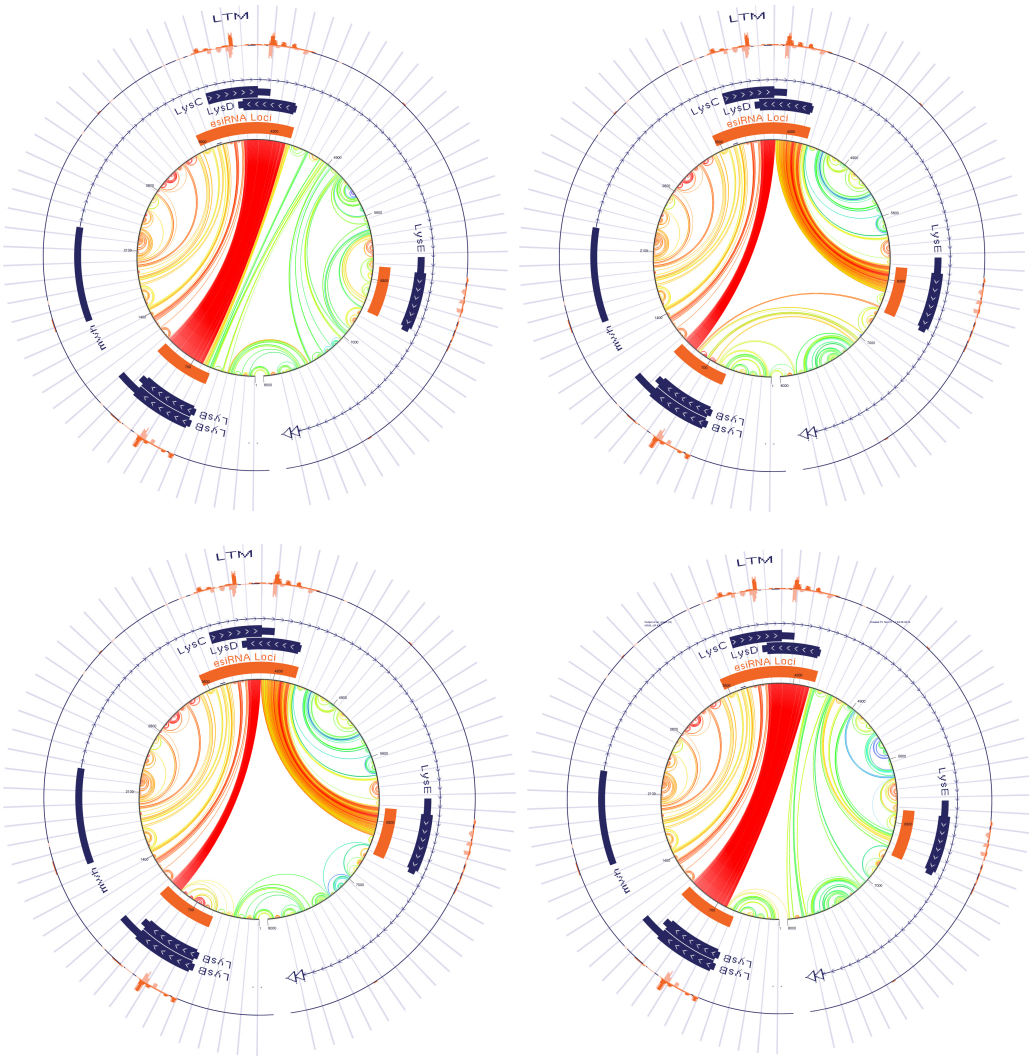


Figure 2.17. (Continued)

4 esiRNA loci display significantly increased read counts in the LTM and Shock only conditions (p= <.00005). These loci also have increased read counts in the odor only condition, but the increase is smaller and not statistically significant. A: All up-regulated esiRNA loci map to lysozyme family genes residing within a single intron of the multiple wing hairs (mwh) gene (indicated by red arrows). B: Most reads mapping to the esiRNA locus within the lysozyme C (LysC) and lysozyme D (LysD) genes fall outside of the region of overlap between the two genes. The regions producing the most reads (highlighted in red and blue) correspond to a long stretch of perfect duplex formed by a hairpin structure in the mwh intron (shown in the lower panel), while the region of overlap between the two genes (highlighted in green), and largely corresponding to the hairpin loop, produces far fewer reads. C: The mwh intron produces several hairpin structures with long stretches of perfectly complementary duplex forming between LysB and LysD (highlighted in yellow and purple, respectively), or LysE and LysD(highlighted in orange and purple, respectively). The lower panel depicts the two predicted secondary structures for the mwh intron with the highest scores, but several other valid predicted secondary structures also form perfect duplexes in the same regions. The red markers denote the extents of RNA sequences depicted in the structures in the lower panel. D: Circle plots showing 4 examples of 50 high confidence structure predictions for an RNA encoded by the 5' 8Kb of the mwh intron. The intron's 5' end is at the 6 o'clock position, and is transcribed clockwise around the circle. Arcs connect base paired nucleotides across the center of the circle plot. Red, orange, and yellow arcs indicate high confidence predictions, while greens and blues indicate lower confidence predictions. Read coverage for the LTM condition is depicted in orange outside of the circle plot, as are the positions of transcripts and esiRNA loci.

Initially, the overlapping antisense arrangement of the LysC/D locus suggested that the reads mapping to the highly similar lysozyme family genes within the mwh intron might all arise from this locus via the cis-NATesiRNA mechanism. However, closer inspection of the LysC/LysD locus unexpectedly reveals that the region of overlap between the two genes produces far fewer reads than do the regions of LysC and LysD outside of the overlap (Fig 2.17 B). We therefore sought to determine if transcription at this locus might yield a hairpin secondary structure that could be a substrate for the esiRNA biogenesis machinery. We used mFold to predict secondary structures for a conceptual RNA whose 5' end corresponds to the transcription start site (TSS) of LysC, and whose 3' end is the LysD TSS.(*60*) The mFold predicted structure for this RNA

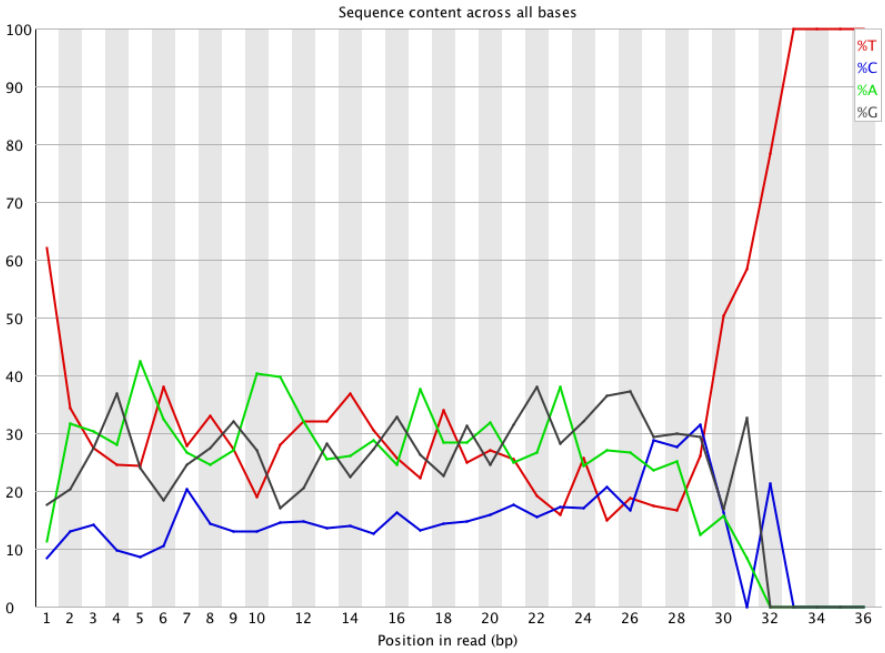
Exhibits a long hairpin structure with a long region of perfect duplex (Fig 2.17 B, lower panel). This duplex region corresponds nearly perfectly to the regions of LysC and LyD that do not overlap, and produce the most reads. The area of reduced read production in the region of overlap between LysC and LysD corresponds to a segment of the secondary structure that contains the hairpin loop, and a number of bulges where no duplex is formed. Thus, it seems likely that reads mapping to the LysC/LysD esiRNA locus are the product of the hairpin esiRNA pathway, and not of cis-NAT driven esiRNA biogenesis. The LysB, LysE and LysS esiRNA loci, also reside within the same mwh intron, and therefore might also correspond to regions of secondary structure duplex formed by the mwh intron. To test this idea, we submitted segments of the mwh intron to mFold. The maximum sequence length capacity of mFold is 9Kb. While LysB, LysC/LysD, and LysE fall within the 5' 7Kb of the mwh intron, LysS is ~15Kb 3' of LysE, which prevented us from testing for secondary structures in which LysS forms duplexes with the other up-regulated esiRNA loci. As the lysozyme family genes closely resemble each other in sequence, we expected that duplexes might form between different combinations of intron sequences corresponding to theses genes. Therefore, we first predicted structures for mwh intron sequences spanning individual pairings of esiRNA loci. Interestingly, while the hairpin formed within the LysC/LysD esiRNA locus did not form extensive duplexes that include the region of overlap between the genes, mwh intron sequences that span the region of LysB through LysD, or LysC through LysE both form long hairpins with extensive duplexes that span the entire length of LysD, including the region of overlap between LysC and LysD (Fig 2.17 C). We next

obtained secondary structure predictions for the 5' 8Kb of the mwh intron using mFold. The structures predicted for this 8Kb sequence all exhibit extensive duplexes between the LysB, and LysC/LysD, and between LysC/LysD and LysE loci, though the particular regions of each locus involved in duplex formation varied between predictions. Lastly, we could not conduct such a secondary structure analysis for the LysS esiRNA locus, as it is too distant from the other esiRNA loci within the mwh intron. However, we note that unlike the other up-regulated lysozyme gene family esiRNA loci, LysS reads are dominated by just two reads. Both dominant LysS reads also map to all other up-regulated lysozyme gene family esiRNA loci. Thus, LysS mapping reads, and their up-regulation may actually reflect esiRNA production from the hairpin structures formed by the 5' end of the mwh intron.

A profile of piRNA expression in the Drosophila head

In a manner similar to that which we used to identify esiRNA loci, we collected piRNA loci from our read-contigs by selecting those read-contigs having at least 250 mapped reads from our combined libraries, with at least 75% of these reads being 24-29nt. This size range corresponds to the known lengths of *Drosophila* esiRNAs. Using these criteria, we identified 82 likely piRNA loci. Reads mapping within piRNA loci display a strong bias for uridine at the 5' nucleotide, and adenosine at the 10th nucleotide, a hallmark of the ping-pong piRNA mechanism (Fig 2.18).

Figure 2.18. Per-base sequence composition of reads mapping to piRNA loci



Reads mapping to piRNA loci were selected using bedtools, and the fraction of reads with each base at each nucleotide position were then plotted using FastQC. piRNA mapping reads have a strong bias for a 5' uridine, and adenosine at position 10, a hallmark of the ping-pong mechanism.

These biases argue in favor of the validity of our designation of these loci as piRNA producing regions. piRNA loci are shorter and have fewer mapped reads than esiRNA loci. 95.06% of the piRNA loci we identified map to repetitive elements (Fig 2.19).

Figure 2.19. Characteristics of piRNA loci

A

Contig Type	Count	Avg. Contig Length	SEM Contig Length
All Contigs	219482	190.74	1.36
All Contigs >= 250 Reads	7581	2176.43	28.83
esiRNA (21-22nt)	368	4536.40	231.73
piRNA (24-29nt)	82	1453.13	275.05

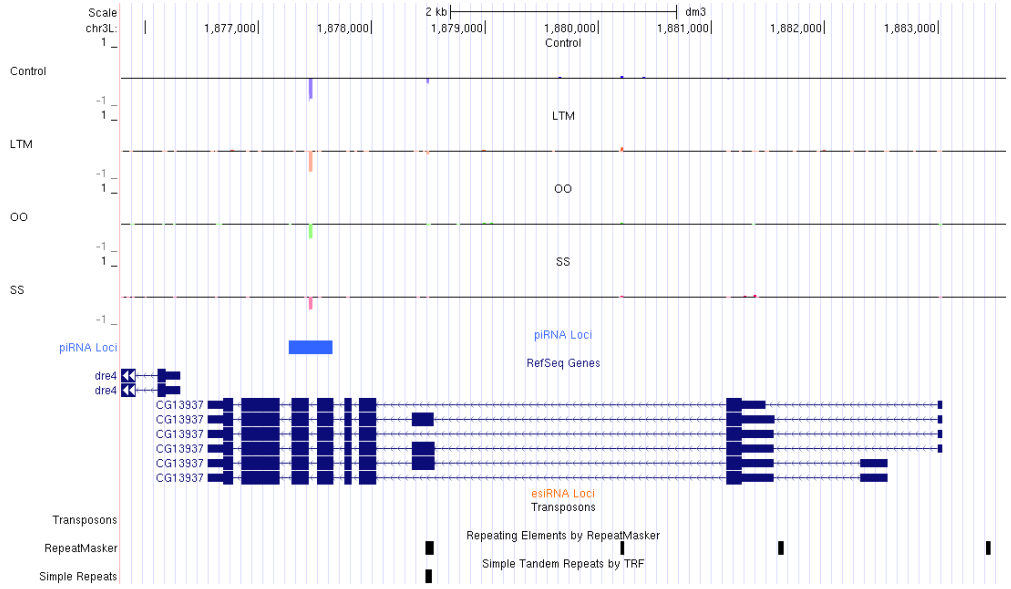
B

	Total	Intron	Exon	5' UTR	3' UTR	Intergenic	Transposon	Gene	Repeat	Gene w/o Transposon	Gene w/o Repeat
# of piRNA loci	82	34	7	10	3	45	16	36	77	33	1
Fraction of all piRNA Loci	100.00%	41.98%	8.64%	12.35%	3.70%	55.56%	19.75%	44.44%	95.06%	40.74%	1.24%

piRNA loci were collected from our read-contigs by selecting those contigs having 75% mapped reads 24-29nt in length. A: mean piRNA locus length is shorter than that for all read-contigs, or esiRNA loci. B: The types of genomic features overlapping each piRNA locus were obtained using bedtools to intersect feature position lists with piRNA loci positions. We then calculated the fraction of piRNA loci overlapping each type of feature.

Surprisingly, more piRNA loci overlap genes than overlap transposons. However, many piRNA loci overlapping genes correspond to repetitive elements that fall entirely within long introns. The one piRNA locus mapping within a gene but not to a repetitive element is dominated by a single 26nt sequence, rather than being comprised of a variety of sequences spread out across the locus as is typically the case. This sequence maps to the 3' end of an intron of the CG13937 gene, and nowhere else in the genome (Fig 2.20).

Figure 2.20. piRNA locus mapping to CG13937 and outside of repetitive elements

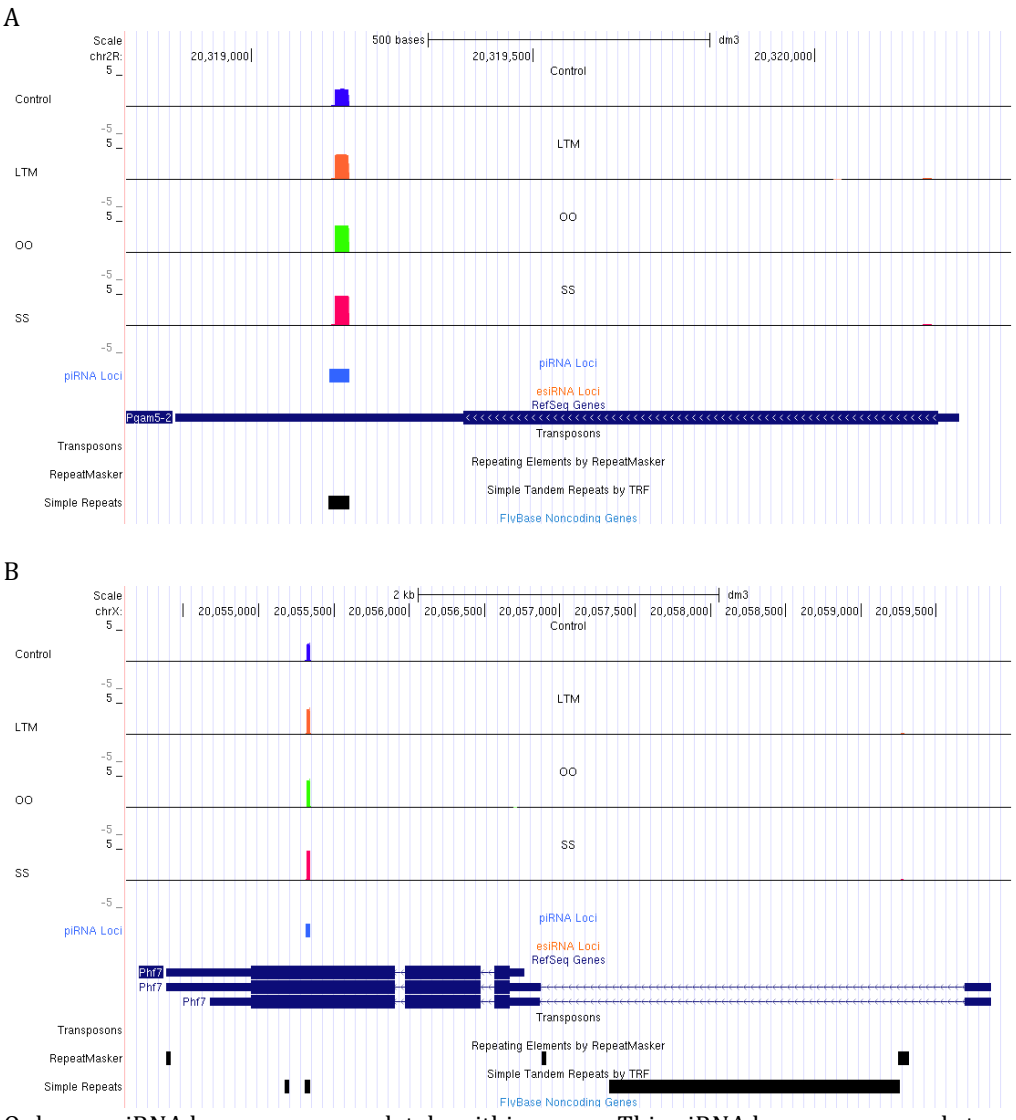


The lone piRNA locus mapping within a gene but outside of repetitive elements is dominated by a single sequence. This sequence maps to the 3' end of an intron of the CG13937 gene.

CG13937 is annotated in FlyBase as a HNK-1 sulfotransferase based upon its amino acid sequence, and is expressed in the larval and adult CNS. However, this gene has no known function. Only two piRNA loci map entirely within exons, both of which correspond to separate instances of the same short repetitive element. Instances of this element occur at several locations across the genome, including in the 3' UTR of

the Phosphoglycerate mutase 5-2 (Pgam5-2) gene, which resides on chromosome arm 2R. The only gene other than Pgam5-2 that this element maps to is the PHD finger protein 7 ortholog (phf7), where it falls within a protein coding exon (Fig 2.21).

Figure 2.21. A repetitive element produces piRNAs antisense to exons



Only one piRNA locus maps completely within exons. This piRNA locus corresponds to a short repetitive element found in several locations throughout the genome. A: The repetitive element generates piRNAs that are antisense to the 3' UTR of the Pgam5-2 gene. B: The same reads mapping to Pgam5-2 also map antisense to a protein coding region within 3' exons of the Phf7 gene.

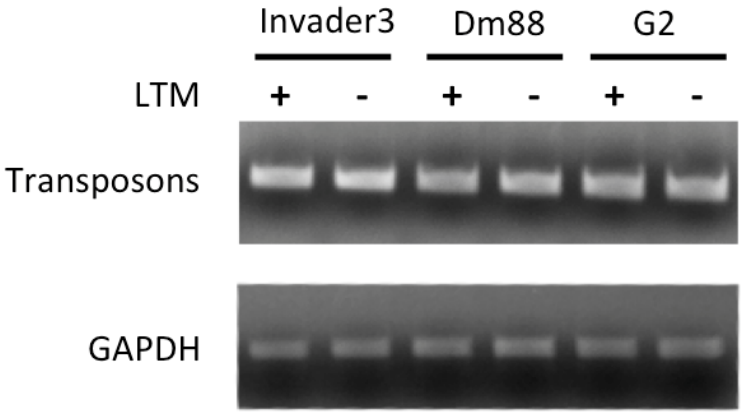
In both instances, piRNAs produced by this short repeat map in the opposite direction of the gene in which the repeat is present.

Changes in piRNA expression during LTM formation

Having identified sites of piRNA production across the genome, we next sought to determine if expression at any of these loci changes during LTM formation. Further, we hoped to resolve any distinct contributions of the US or CS to such changes. We used edgeR in its GLM mode to test for statistically significant differential expression between our experimental conditions. EdgeR analysis identifies 10 piRNA loci exhibiting significant changes in expression during in the LTM condition, but no changes in the shock only or odor only conditions are significant. When compared to control, all 10 of these loci have decreased read counts in the LTM condition. Further, all 10 significantly down-regulated piRNA loci correspond to intergenic repetitive elements. 2 transposons, Dm88 and Invader3, are present in the repetitive elements exhibiting significantly decreased read counts. Both belong to the long terminal repeat (LTR) class of retrotransposons. As such, decreased piRNA expression at these loci might result in increased expression of their corresponding transposons. As our libraries were prepared from size selected sRNAs, we used semi-quantitative RT-PCR to examine expression of these transposons in the same total RNA samples from which our sequencing libraries were prepared. We used random hexamers to prime the reverse transcription reaction, and primers directed at sequences in the Invader3 and Dm88 LTR

retrotransposons, as well as the non-LTR transposon G2 to amplify the cDNA. Primers directed at a GAPDH sequence were used as an internal control. In 3 replicates, expression of all 3 transposons examined remained essentially unchanged (Fig. 2.22).

Figure 2.22. Examination of transposon expression using semi-quantitative RT-PCR



Semi-quantitative RT-PCR was used to examine expression of the LTR retrotransposons Invader3 and Dm88, as well as the non-LTR transposon G2 during LTM Formation. GAPDH was used as an internal control for all reactions. In all cases, transposon expression remained essentially unchanged.

Discussion

Formation of lasting memories is known to involve synaptic plasticity mediated by changes in transcriptional programs and protein synthesis. sRNAs influence both of these mechanisms in a number of ways. A substantial and growing body of evidence demonstrates the critical importance of translational control via the microRNA pathway in regulating synaptic plasticity, and in memory formation. Much of what is known about the involvement of microRNAs in memory formation comes from studies of individual microRNAs. However, the many microRNAs

expressed in a single cell engage in complex and overlapping regulatory cascades. In many cases, target sites for multiple microRNAs are present within a single mRNA. Thus, a comprehensive understanding of the changes in gene expression driven by microRNAs during memory formation requires simultaneous observation of all microRNAs present in the brain. To our knowledge, this work is the first genome wide examination of microRNA expression during memory formation in *Drosophila*.

The profile of microRNA expression in the fly head that we obtained differs somewhat from that of Berezikov et al, which is perhaps the most comprehensive examination of microRNA expression in *Drosophila* to date.(*35*) In both cases, the most highly expressed mature microRNAs represent a large fraction of all microRNA reads. However, we find that miR-184-3p is the most highly expressed mature microRNA, accounting for ~36% of all raw microRNA reads, while Berezikov et al find that miR-184-3p is only the 30th most highly expressed mature microRNA, representing just ~0.3% of all raw mature microRNA reads. Similarly, Berezikov et al find that let-7 is the most highly expressed mature microRNA in head lysates, comprising ~25% of raw mature microRNA reads, while it is the 7th most abundant mature microRNA read in our libraries, representing just ~3% of such reads. In both studies, the 10 most highly expressed microRNAs constitute the overwhelming majority of mature microRNA sequences, representing ~90% and ~75% of mature microRNA reads in this study and in Berezikov et al., respectively. It is likely that both data sets suffer from cloning artifacts that could have been introduced at the

adapter ligation or PCR amplification steps. Such artifacts are known to be produced by the Illumina sRNA prep kit v1.5, which both studies utilized. Further complicating comparison of the two studies, the Berezikov et al data includes several libraries produced by the authors themselves, as well as several previously published head specific data sets. The libraries produced by the authors were prepared and sequenced similarly to our libraries, but those obtained from previous work were more varied, including libraries sequenced using 454 technology and others that had undergone beta-elimination. Thus, the overrepresentation of the most highly expressed microRNAs in the Berezikov et al data may be somewhat reduced by this diversity of source material, while the more uniform libraries in our study likely exacerbated such artifacts. However, when comparing mature microRNA expression between the libraries used in this study, we find that expression ranks are largely consistent. This suggests that if overrepresentation of certain microRNAs in our libraries is an artifact, it is systematic. Berezikov et al do not provide per-library counts, and it is therefore not clear how consistent microRNA expression is across the libraries they used.

In addition to profiling miRNA expression in the *Drosophila* head, we sought to identify changes in this profile resulting from LTM formation. Further, our use of the well characterized aversive olfactory classical conditioning paradigm allowed us to examine changes in microRNA expression uniquely induced by the CS or US, as well as those that result from the combination of CS and US during conditioning.

Again, this work is, to our knowledge, the first genome-wide examination of changes in microRNA expression to make such a distinction.

Using this approach, we show that miR-312-3p is the lone mature microRNA significantly down-regulated during LTM formation. While miR-312-3p expression is also reduced in the odor only and shock only conditions, our analysis did not find these changes to be statistically significant. We had 6 samples in the control condition, 5 samples in the LTM condition, but just 3 samples each in the shock only and odor only conditions. Thus, the lack of statistical significance in the odor only and shock only conditions may simply reflect the reduced power of our analysis in these groups. The reduction in miR-312-3p expression is more pronounced in the LTM and shock only conditions than in the odor only condition (-1.66, -1.32, and -0.71 log fold change in concentration respectively). These findings suggest that reduced miR-312-3p expression in the drosophila brain during LTM formation results largely from signaling via the US circuitry, which involves dopaminergic signaling in the MB. We examined the predicted targets of miR-312-3p in the hope of identifying candidate genes for future study. Interestingly, miR-312-3p targets the dopamine receptor Dop1R1. Previous work demonstrates that both microRNA and dopamine receptor expression in the corresponding areas of the brain are required for LTM in *Drosophila*.(*42*) miR-312-3p also targets the octopamine receptor Octbeta3R. Octopamine signaling is known to be activated in the *Drosophila* brain during conditioning.(*61*) Additionally, a recent publication by Wu et al demonstrates that octopaminergic circuitry in the *Drosophila* MB is involved in

ARM formation.(*62*) Lastly, miR-312-3p negatively regulates synaptic efficacy at the NMJ via silencing of kinesin heavy chain expression.(*21*) Thus, our finding that miR-312-3p is significantly down-regulated during LTM formation fits with a model in which LTM relevant neuronal activity increases synaptic efficacy in appropriate circuits in part by relieving key pathways from miR-312-3p mediated silencing. Additional experiments will be required to determine the validity of such a model, and to identify the neurons involved. We also identified 4 mature microRNAs up-regulated during LTM formation. miR-314-3p, miR-956-3p, miR-958-3p, and miR-958-5p all display statistically significant up-regulation in the LTM condition. Only miR-958-5p is significantly up-regulated in the shock only condition. No change in the expression of any microRNA is statistically significant in the odor only condition. However, of the microRNAs significantly up-regulated in the LTM condition, miR-958-5p increases the most in all three experimental conditions. Observed changes in expression of microRNAs in the odor only condition are generally smaller than in the shock only or LTM conditions. This may simply reflect the activation of a smaller set of neurons by odor than by shock. However, it may also reflect the nature of signaling activated, as the US is known to trigger activity in the dopaminergic circuitry of the MB. We generated our libraries from whole head lysates, and we are therefore unable to determine which cell types contribute to observed changes in microRNA expression. It is possible that a small number of cells are exhibiting large changes, or that a larger set of cells undergoes more modest changes. Further, we are unable to determine if all changes in microRNA expression occur in the same cells, or whether different cell types are responsible

for each significantly altered microRNA. However, previous studies suggest that alterations in neural activity in a single cell type can induce changes in the expression of many microRNAs simultaneously.(*38*) Nonetheless, results from this study point the way for future investigations using other methods better able to resolve such distinctions. Further, our data strongly suggest a number of genes for further investigation in the context of microRNA activity during memory formation. A number of mRNAs are targeted by more than one of the mature microRNAs that are significantly regulated in the LTM condition. mRNAs of 10 genes are targeted 3 or more of the 5 significantly regulated microRNAs that we identified. This set includes Dop1R1, the small conductance calcium activated potassium channel (SK), and No receptor potential A (NorpA), all of which directly modulate neural activity. Intriguingly, the dopamine receptor Dop1R1 is targeted by 3 out of 4 upregulated microRNAs as well as by miR-312-3p. This may reflect the importance of post-transcriptional regulation of Dop1R1 in memory formation. While it may seem contradictory that several microRNAs whose expression changes with the opposite sign following LTM training would target the same gene, the combined regulatory output of these microRNAs may work synergistically to facilitate olfactory memory. It may be that the up- and down-regulated microRNAs are expressed in different sets of cells. One could conceive that neurons involved in the memory trace might need enhanced sensitivity to dopamine, while those that are not require reduced dopamine sensitivity. Competitive regulation of Dop1R1 by these 4 microRNAs in the same cells might also have important effects. It is also important to note that, while Dop1R1 is targeted by 4 regulated microRNAs, only dme-mir-312-3p targets

Dop1R2 and Octβ3R. Thus, the combined effect of dme-mir-312-3p and the 3 upregulated microRNAs that also target Dop1R1 may be to shift dopamine reception to Dop1R2 and to enhance sensitivity to octopamine. Determining which, if any, of these models is correct will require further experimentation. A key step will be determining the expression patterns of regulated microRNAs within the brain. Whether or not these microRNAs are expressed in the same cells will obviously affect the regulatory consequences of changes in their expression levels. Another key set of experiments will be to examine how response to dopamine is modulated by expression of these microRNAs, and to determine if such an effect is important for memory formation.

To better understand which pathways are subject to regulation by microRNAs whose expression change during memory formation, we conducted a gene ontology analysis on the sets of genes targeted by these microRNAs. We find that genes involved in G protein coupled signaling are overrepresented in this set. G protein coupled signaling regulates synaptic activity and efficacy in important ways, most directly through a variety of G protein coupled receptors present at synapses. MicroRNA mediated control of local translation could provide a way by which G protein coupled signaling is selectively tuned for subsets of the many synaptic connections in which a given neuron engages. Such selective alterations in neural circuits are integral to models of memory formation, but the molecular mechanisms involved are not fully understood. We also identified a number of genes annotated with the term “neurological systems process” that are targeted by more than one of

the microRNAs significantly regulated during LTM formation. Genes in these sets are obvious avenues for an in depth examination of the function of microRNAs in regulating synaptic plasticity and memory.

Though microRNA sequences are dominated by a few canonical mature microRNAs, a diversity of non-canonical microRNAs are present in cells.(*33*) Previous studies of microRNA involvement in *Drosophila* synaptic plasticity and memory have exclusively focused on canonical mature microRNAs. Our use of massively parallel sequencing allowed us to examine our reads for signs of changes in microRNA processing across our experimental conditions. Such changes might provide a mechanism through which microRNA activity could be regulated. We observed numerous instances of non-canonical 3' end formation for several microRNAs, but very few instances of non-canonical 5' ends for all but a few miRNA. These results largely resemble those reported for a similar analysis by Berezikov et al.(*35*) For instance, we observe that ~46% of miR-79-3p reads are 5' offset, while Berezikov et al report the rate to be ~30%. Some differences do exist between our results and those reported by Berezekov et al, but their analysis of non-canonical microRNA ends pooled reads from many tissue types, cell lines, and mutant strains.(*35*) The substantial similarity between the results of our non-canonical microRNA end analysis and those of Berezikov et al therefore supports the view that for most microRNAs, 5' end precision is strictly maintained. As the 5' seed sequence is the major determinant of microRNA target specificity, 5' precision is vital for proper microRNA function. The overwhelming precision of 5' ends we observed

shows that imprecise 3' cleavage, and not offset microRNA processing, is the major cause of non-canonical 3' ends present in our data. Though we document extensive 3' imprecision, we observe no statistically significant changes in microRNA processing between our experimental conditions. Massively parallel sequencing also permitted investigation of untemplated nucleotide addition to microRNA reads. Such modification is thought to reflect the activity of mechanisms that regulate the stability of microRNAs.(*49*, *63*) We find that 5' tailing is extremely rare, and that 3' tailed reads represent a small fraction of microRNA reads, with only miR-927-3p having greater than 5% of its reads tailed. miR-927-3p tailing events are almost exclusively instances of mono- and poly-adenylation. miR-927-3p is the minor species produced from the mir-927 hairpin. For microRNAs expressed at high enough levels in both our work and head libraries from Berezikov et al, we find similar rates of 5' and 3' tailing.(*35*) Though we were able to observe 3' tailing of microRNAs, we did not observe any statistically significant changes in 3' tailing between our treatment conditions. Our analytical techniques exclude instances of 3' modification extending beyond 2nt. However, previous studies have demonstrated that microRNAs with untemplated extensions beyond 2nt also have a substantial number of 2nt tailing events.(*45*, *64*) Thus, we believe that our methods allow us to obtain a representative view of 3' tailing events.

As massively parallel sequencing allows observation of untemplated nucleotides, we were able to search for instances of microRNA editing via ADAR activity. ADAR is known to influence expression of mRNAs in the brain, and

influences neuronal activity and behavior in *Drosophila*.(*65*, *66*) Furthermore, modification of microRNA precursors by ADAR regulates expression of polycistronic microRNAs.(*67*) We observed what appears to be genuine A to I editing events in 7 microRNAs. Both our analysis and berezikov et al. identify a candidate editing event in miR-971.(*35*) However, the other 6 candidate editing events we identify and 3 candidate editing events identified by Berezikov et al in wild type head libraries do not overlap. This difference is likely do at least in part to our exclusive use of Canton-S flies, and the inclusion of libraries from Oregon-R fly heads in their analysis. Several of the candidate editing events we identify occur at positions within the mature microRNA sequence, and are therefore likely to influence the silencing activity of these microRNAs. However, published validated microRNA target predictions are based solely upon the canonical microRNA sequence. As de-novo target prediction or validation is beyond the scope of this work, we are unable to comment on the potential downstream regulatory consequences of these editing events. Instances in which editing falls outside of the mature microRNA sequence may lead to altered expression of mature microRNAs, but again we are unable to predict the outcome of such editing. None of the microRNA editing events we observed exhibited statistically significant differences between treatment groups. Thus, our data does not indicate that RNA editing is a major mechanism regulating microRNA activity during memory formation.

Taken together, our profile of microRNA sequences present in the *Drosophila* head, and characterization of the changes in these sequences resulting from

aversive olfactory classical conditioning, represent the most comprehensive view of microRNA response to long term memory formation in the insect brain to date. Our data provide clear direction for future work towards an understanding of microRNA involvement in memory. We recognize that our use of whole head lysates likely obscures subtle changes in many neurons, and perhaps significant changes in a few neurons. However, our inclusion of all cell types present in the brain at the least provides the opportunity to detect such changes, if they do occur. While we cannot rule out such possibilities, the statistically significant changes in microRNA expression that we were able to detect likely represent key regulatory events, and should be pursued further. To our knowledge, this study is the first detailed examination of non-canonical microRNA sequences in the context of insect memory formation. Though we did not identify any statistically significant changes in the classes of non-canonical sequences we examined in any of our treatment groups, the regulatory implications of observed microRNA editing events and other modifications are unknown, and may yet yield new insight.

Though others have previously profiled expression of esiRNAs or piRNAs in the *Drosophila* head, this study is, to our knowledge, the first examination of the expression of these classes of sRNAs during memory formation in insects. We identified a set of putative esiRNA loci and profiled sRNA expression at these loci in each of our treatment groups. Preliminary examination of our data revealed a substantial number of what appeared to be esiRNA reads outside of loci identified in previous studies. One such example occurs within the tkv esiRNA producing gene.

Czech et al report a large esiRNA peak corresponding to a region of overlap between tkv and the pseudogene CG14033, which are located on opposite strands.(*57*) Though we too observe 21nt reads in this region of overlap, we do not see a large peak as reported by Czech et al. However, we observe a large number of 21nt reads mapping to both strands at the 3' end of the tkv gene, which is in a cis-NAT arrangement with the TpnC25D gene. Such a signal is the characteristic signature of a cis-NAT esiRNA locus. Czech et al do not report a peak in reads at this cis-NAT region. This discrepancy is likely due to our use of head lysates in this study and the use of testis by Czech et al. The paucity of reads we observe at the CG14033 overlap may result from preferential use of the tkv-B and tkv-D isoforms in heads, whose transcription start sites are 3' of the tkv sequence overlapping CG14033, and therefore do not yield transcripts complementary to CG14033. Alternatively, the CG14033 pseudogene may not be strongly coexpressed with tkv in the head. Similarly, the absence of a peak at the region of overlap between tkv and TpnC25D in the Czech et al data may result from a lack of tkv/TpnC25D cotranscription in the testis.

Many studies reporting the genomic locations of esiRNA production use RNAs extracted from gonads, as *Drosophila* esiRNAs were first recognized in studies of sRNAs expressed in these tissues. We reasoned that reliance on previously reported esiRNA loci might therefore exclude potentially interesting head specific loci, or loci only revealed by neural activity induced during classical conditioning. Accordingly, we performed de novo esiRNA locus identification. Our identification

of several previously unreported esiRNA loci within lysozyme family genes prove the value of this step. Our differential expression analysis shows that the lysozyme family loci are the only esiRNA producing regions mapping at or near genes that whose expression changes significantly during LTM formation. While reads are present at these loci in the control condition, read counts increase dramatically in the LTM and shock only conditions. These changes are highly significant ($p<=0.0005$). Read counts at these loci also increase in the odor only condition, but this increase is smaller, and not statistically significant. Differential expression analysis shows that, though read counts are higher in the shock only condition than in LTM, the difference between LTM and shock only is not statistically significant. This indicates that neural activity driven by the US (shock) is responsible for the observed increase in lysozyme family esiRNA expression. This circuitry has been characterized in the *Drosophila* brain, and is largely dopaminergic. Accordingly, this population of cells is an obvious target for investigation into the significance of lysozyme family esiRNAs.

LysC and LysD are in a cis-NAT arrangement, and all lysozyme family genes exhibiting increased read counts in the LTM and shock only conditions share stretches of sequence identity. One might therefore expect that reads mapping to these lysozyme family genes are the product of LysC and LysD cotranscription. However, careful examination of read coverage across the LysC/LysD locus reveals an unusual signature in which read coverage is dramatically higher in the non-overlapping segments of these genes than in the region of overlap. Long inverted

repeats that form RNA hairpins are known to produce functional esiRNAs in the region of perfect duplex, but not from the hairpin loop.(*59*) Such an arrangement might better explain the pattern of reads at the LysC/LysD locus than a cis-NAT mechanism. To explore such a possibility, we predicted secondary structures for several conceptual RNAs spanning this locus. In all cases, the RNAs formed extensive stretches of perfect duplex at sequences corresponding to the areas of the LysC/LysD locus producing the most reads. Further, the overlapping portion of the LysC/LysD locus, which produces few reads, corresponded to the hairpin loop and adjacent bulged duplexes. These observations argue strongly that esiRNA like reads at the LysC/LysD locus are produced via the hairpin pathway, and not by a cis-NAT mechanism. Most of the reads mapping to LysB and LysE also map within the LysC/LysD locus. LysB and LysE reads could primarily be of LysC/LysD origin. We noted that the pattern of coverage peaks across LysB and LysE also resemble the pattern of peaks across the LysC/LysD locus. However, the patterns of peaks do not match closely enough to be explained simply by multi-mapping reads. This observation raised the intriguing possibility that a much longer progenitor RNA spanning the LysB – LysE region might form duplexes that yield esiRNAs via the hairpin mechanism. Indeed, Mfold predicts hairpin structures spanning this region, with duplexes between LysC/LysD and LysB or LysE. LysC/LysD, LysB-LysC/LysD, and LysE-LysC/LysD duplexes are mutually exclusive, as these structures involve overlapping segments of LysD. No single structure out of those just discussed is capable of producing the observed pattern of read coverage across the LysB-LysE region. Thus, It is likely that esiRNA production from this region involves a

combination of several hairpin structures. LysS is located too far away from the other lysozyme family genes in this region to permit structure prediction using Mfold. However, we note that the coverage peak mapping within LysS corresponds to a handful of unique read sequences that are also present in the LysB-LysE region. It is therefore likely that reads mapping to LysS are actually the result of esiRNA production from LysB-LysE hairpins, and not from LysS itself. The regulatory function of esiRNA production from the LysB-LysE region is unclear. esiRNAs produced from this region should downregulate LysB, LysC, LysD, LysE and LysS. However, lysozymes primarily function in defense against bacterial infection by catalyzing the hydrolysis of peptidoglycans present in bacterial cell walls. As such, the reason that reduced Lysozyme expression would be needed during LTM formation is not obvious. An alternate explanation for increased LysB-LysE esiRNA production during LTM formation could be that the increase reflects increased expression of the mwh-B or mwh-C isoforms. These isoforms both include the intron that spans the full LysB-LysS region, while the TSS for the mwh-A isoform is downstream of LysB, and therefore does not include this intron. Mwh is a G-protein binding domain-formin homology 3 (GBD-FH3) protein in the frizzled pathway that negatively regulates actin polymerization and filament formation, and is a planar cell polarity (PCP) effector.(*68*, *69*) Thus, one possible interpretation of this result is that increased production of LysB-LysE esiRNAs reflects regulation of mwh in the course of reorganizing the cytoskeleton to accommodate altered neural connectivity. Mwh has been studied largely in the context of PCP during wing hair formation, and no role for mwh in synaptic plasticity has yet been demonstrated.

However, as a regulator of the cytoskeleton, mwh is a more obvious target for regulation in the course of memory formation than are lysozyme family genes. This possibility should be further investigated, as it would be the first instance of an esiRNA mediated change in gene regulation in response to memory formation documented in insects. Further, such investigation would require exploration of the possibility that lysozyme family genes are regulated during memory formation. While a negative result would support mwh involvement in synaptic plasticity, a positive result would necessitate a reexamination of the functions of lysozyme genes in the *Drosophila* head. Both possibilities provide the potential for novel insights into the basic mechanisms of memory.

Our methods also allowed us to identify genomic regions that produce sequencing reads with characteristics typical of piRNAs, including length distribution and a 5' uridine bias. However, our study did not include methods such as β-elimination that would more definitively categorize these reads as genuine piRNAs. Nonetheless, the fact that many of our piRNA loci correspond to those identified in studies that did include these steps gives us confidence that our methods yield meaningful piRNA profiles.(*33*, *54*) We observe only subtle changes in piRNA profiles in the LTM treatment group, and insignificant changes in the odor only or shock only conditions. Those piRNA loci exhibiting statistically significant changes in the LTM condition exclusively mapped to repeats and LTR retrotransposons. In all such instances, reads decreased in the LTM condition vs. control. None of these loci produce reads that map perfectly to mRNAs or to known

regulatory elements. Furthermore, semi-quantitative RT-PCR did not detect significant changes in the levels of transcripts corresponding to these LTR retrotransposons. Thus, there are no obvious consequences for gene or transposon regulation stemming from changes in piRNA expression at these loci. We did however identify piRNA loci with interesting features, perhaps warranting further investigation. In one instance, a 26nt read uniquely maps to the 3' end of an intron of the CG13937 gene. This read does not correspond to a repeat or transposon, and is therefore atypical for a piRNA, if it is in fact loaded into piwi proteins. The 3' end of the read maps precisely to the 5' end of a CG13937 exon, suggesting that this might be a miRtron. However, at 26nt, this read is longer than would be typical for a microRNA, and this sequence is not found in the miRbase catalog of *Drosophila* microRNAs, which includes miRtrons. Again, it is difficult to speculate on what regulatory significance this sequence has. Another interesting piRNA locus corresponds to a short repeat found within exons of the Pgam5-2 and Phf7 genes. In both cases, the read is antisense to the overlapping gene. While no statistically significant change in the expression of this read is observed in any of our treatment conditions, the unique characteristics of this locus may warrant further study. Rajasethupathy et al. show that piRNAs are expressed in *Aplysia* neurons, and are regulated by memory relevant neurotransmitters.(*5*) Serotonin exposures that induce LTP result in downregulation of the transcriptional repressor CREB2. This occurs via piwi and DNA methyltransferase (DNMT) dependent methylation of the CREB2 promoter. Antisense inhibitors of piRNAs mapping at the translational start site of CREB2 relieve this repression. Thus, their results support the idea that

piRNAs are involved in memory relevant gene regulation, and lead us to search for evidence of such a mechanism in our data. However, differences exist in the piRNA pathways of *Aplysia* and *Drosophila.* Furthermore, DNA methylation in *Drosophila* appears to occur at different locations, and via a mechanism that differs substantially from those of other model organisms. *Drosophila* do not possess a DNMT-1 homologue, and DNMT-2 is dispensable for methylation of their DNA.(*70*) Thus, it is possible that no mechanism analogous to piwi/piRNA mediated CREB2 regulation in *Aplysia* exists in *Drosophila*. piRNA function in the fly brain remains poorly studied, and may yet prove relevant to memory. Though our methods failed to detect changes in expression of piRNAs that would have obvious implications for gene regulation, they may have limited sensitivity. Therefore, our results do not preclude the possibility that piRNAs mediate aspects of memory formation in *Drosophila*.

Expanding our examination of sRNA expression in the fly head beyond microRNAs to include esiRNAs and piRNAs allowed us the first view of how these classes of sRNAs respond during memory formation in insects. Though we observe few statistically significant changes, our data hint at previously unidentified regulatory functions for these sRNAs, and at novel ways in which gene regulation is altered by neural activity and memory formation. The extensive conservation of the esiRNA and piRNA pathways indicates that both have vital functions in animal biology. Though much has been learned about their roles in defense of the genome, a great deal remains to be understood about how they influence gene expression.

Further, much of what is known about esiRNAs and piRNAs comes from studies in *Drosophila* gonads, tissues with unique properties that may prohibit the generalization of findings from these studies to cell types from other organs. As such, a great deal of work remains to be done before these sRNA classes are fully understood. Due to its central role in the discovery of esiRNAs and piRNAs, and the available repositories of sequencing data produced in those efforts, *Drosophila* remains one of the best model organisms in which to study these classes of sRNAs. This study is one of the first attempts to examine how esiRNA and piRNA expression changes in response to the demands of altered physiology in wild type animals. The genetics of learning, memory, and behavior are perhaps better understood in *Drosophila* than in any other model organism. Accordingly, the fly brain will likely prove to be a productive system in attempts to identify the possible functions of esiRNAs or piRNAs in regulating gene expression in neurons. Results from this study support such a view.

Defects in posttranscriptional control of gene expression are emerging as the underlying causes of many diseases of the brain and nervous system. sRNAs are central players in mechanisms that govern the stability, localization, and translation of mRNAs. Numerous cases of individual microRNAs required for normal neural function, memory, and behavior have been identified in many model organisms, and in humans. Yet we know that microRNAs do not operate in isolation. As such, a better understanding of the functions of microRNAs in these processes will require genome wide approaches. Further, the roles that esiRNAs and piRNAs play in

neurons remain poorly understood. As few cases of individual genes strongly regulated by these sRNA classes have been identified, let alone well characterized, esiRNAs and piRNAs may still be best studied through genome wide surveys of their expression in various experimental settings. This study is an early attempt to address these needs. Future work with a tighter focus on specific cell types, sRNA/target pairings, or other refinements may better inform our understanding of sRNA biology in neurons, and in memory formation more broadly. Indeed, as this study does not include experiments examining sRNA/target interactions, or behavioral studies in which the regulated sRNAs we identify are misexpressed, we can not draw direct conclusions about the relevance to LTM of the changes in sRNA expression we observe following conditioning. However, this study does provide foundational information, based upon which such experiments can be designed.

Materials and Methods

Drosophila rearing

CS-Quinn flies were reared on standard cornmeal medium at 25 °C under a 12hr light/dark cycle. 1 Day old adults were used in all experiments.

Training and memory assay

Training was conducted using a semi automated conditioning apparatus as described in (*40*). Briefly, a single batch of flies was split into one of four groups: Control, odor only, shock only, or shock + odor (LTM), and housed in bottles containing no food for one hour prior to training. Flies were then loaded into vials that permit the flow of air from the apparatus and contain electrode grids. Odor and shock delivery are controlled by a computer. Following training, flies were returned to bottles containing no food, and allowed to rest for 2 hours before dissection. Memory formation in LTM flies was evaluated using a T-maze apparatus as described in (*13*). The performance index (PI) was calculated as the number of flies avoiding the conditioned odor minus the number of flies avoiding a control odor divided by the total number of flies in the experiment.

Tissue processing and total RNA extraction

For each sample, heads of 10 male and 10 female flies were dissected and immediately flash frozen and stored at -80 °C while awaiting results from memory testing. Only samples from matched batches of flies in which the LTM trained flies

had a PI >=20 were used in subsequent steps. Heads were homogenized in Qiazol (Qiagen) using a motorized tissue homogenizer. Total RNA was extracted using the Qiagen miRNEasy micro kit, according to the manufacturer's protocol. RNA concentration was obtained using a Nanodrop and Bioanalyzer. RNA quality was determined by Bioanalyzer analysis.

sRNA sequencing

Prior to sequencing library preparation, total RNA used for sRNA sequencing was size selected by PAGE, as recommended in the Illumina sRNA sample prep kit (v1.5) protocol. Gel slices containing only 15-35nt RNAs were excised and used in downstream library preparation steps. sRNA sequencing libraries were prepared using the Illumina sRNA sample prep kit (v1.5) according to the manufacturer's protocol. Library quality and concentration were determined by Bioanalyzer analysis. sRNA sequencing was conducted by Harvard systems biology core facility staff, using an Illumina Genome Analyzer IIx.

Sequencing data analysis

Adapter sequences were removed using the FASTX toolkit. For microRNA analysis, sequencing reads were aligned to miRBase release 19 Drosophila hairpin sequences using Bowtie v0.12.7.(*71*, *72*) Custom software written in the R and AWK languages were used for canonical and isomiR analysis. For esiRNA and piRNA analysis, reads were aligned to the FlyBase Drosophila melanogaster genome (v5.48) using Bowtie v0.12.7.(*71*). SAMtools was used to convert alignments to BAM format.(*73*)

BedTools v2.18.2 was used to count reads aligning within a given feature, and to merge reads into read-contigs.(*74*) The UCSC genome browser was used to visualize sequencing data.(*75*)

Differential expression analysis

The Bioconductor package edgeR v3.1.9 was used in its GLM mode for read count differential expression analysis. Read counts were normalized using the upper quartile method.(*76*) For isomiR analysis, the percentage of normalized reads with a given non-canonical feature was computed for each miRNA. Two tailed students t-tests were used to test for statistical significance of differences in these percentages. P-values were then adjusted for multiple testing using the Holm-Bonferroni method.

Target prediction

DIANA microT v4 predictions were used for microRNA target analysis.(*24*)

Gene ontology analysis

The PANTHER gene ontology (GO) database was used to assign GO terms and search for over/underrepresented terms.(*44*)

RNA secondary structure prediction

Mfold was used to predict RNA secondary structures.(*60*)

Semiquantitative RT-PCR

First strand cDNA was reverse transcribed from 500ng input total RNA, using random hexamer primers (Invitrogen) and the Superscript II kit (Invitrogen). Reverse transcription occurred at 42 °C for 1 hour. Oligonucleotide pairs spanning transposon sequences were used to prime 20 cycles of PCR amplification. Perfect Taq (5 prime) polymerase was used for PCR amplification.

Primer sequences:

Gapdh-FWD: AGCTGATCTCTTGGTACGACAAC
Gapdh-REV: ATGCTTATGAGTCGGCATTTTTA

Inv3-FWD: CCTTTAGCCAACTTCACGACGG
Inv3-REV: GGAATTCGAATTGCCCTAACGG

G2-FWD: GGCAATCAAAACTCTCACGGATG
G2-REV: GGGGATTTGCTAGCCTTTAGG

DM88-FWD: GGATACTCTGATGCTTCTAAGGG
DM88-REV: CACTGCAAAGACCCATTTTGAC

Literature Cited

1. C. H. Bailey, E. R. Kandel, K. Si, The Persistence of Long-Term Memory. *Neuron* **44**, 49–57 (2004).

2. M. Mayford, S. A. Siegelbaum, E. R. Kandel, Synapses and Memory Storage. *Cold Spring Harb Perspect Biol* **4**, a005751–a005751 (2012).

3. B.-T. Juang *et al.*, Endogenous Nuclear RNAi Mediates Behavioral Adaptation to Odor. *Cell* **154**, 1010–1022 (2013).

4. E. McNeill, D. Van Vactor, MicroRNAs Shape the Neuronal Landscape. *Neuron* **75**, 363–379 (2012).

5. P. Rajasethupathy *et al.*, A Role for Neuronal piRNAs in the Epigenetic Control of Memory-Related Synaptic Plasticity. *Cell* **149**, 693–707 (2012).

6. P. Rajasethupathy *et al.*, Characterization of Small RNAs in Aplysia Reveals a Role for miR-124 in Constraining Synaptic Plasticity through CREB. *Neuron*

63, 803–817 (2009).

7. R. Fiore *et al.*, Mef2-mediated transcription of the miR379-410 cluster regulates activity-dependent dendritogenesis by fine-tuning Pumilio2 protein levels. *The EMBO Journal* **28**, 697–710 (2009).

8. R. L. Davis, Traces of Drosophila memory. *Neuron* **70**, 8–19 (2011).

9. C. Margulies, T. Tully, J. Dubnau, Deconstructing memory in Drosophila. *CURBIO* **15**, R700–13 (2005).

10. J. L. Pitman *et al.*, There are many ways to train a fly. *Fly (Austin)* **3**, 3–9 (2009).

11. T. Tully, Drosophila learning: behavior and biochemistry. *Behav. Genet.* **14**, 527–557 (1984).

12. W. G. Quinn, W. A. Harris, S. Benzer, Conditioned behavior in Drosophila melanogaster. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 708–712 (1974).

13. T. Tully, W. G. Quinn, Classical conditioning and retention in normal and mutant Drosophila melanogaster. *J. Comp. Physiol. A* **157**, 263–277 (1985).

14. Y. Dudai, Y. N. Jan, D. Byers, W. G. Quinn, S. Benzer, dunce, a mutant of Drosophila deficient in learning. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1684–1688 (1976).

15. W. G. Quinn, P. P. Sziber, R. Booker, The Drosophila memory mutant amnesiac. *Nature* **277**, 212–214 (1979).

16. K. W. Choi, R. F. Smith, R. M. Buratowski, W. G. Quinn, Deficient protein kinase C activity in turnip, a Drosophila learning mutant. *Journal of Biological Chemistry* **266**, 15999–15606 (1991).

17. T. Tully, T. Preat, S. C. Boynton, M. Del Vecchio, Genetic dissection of consolidated memory in Drosophila. *Cell* **79**, 35–47 (1994).

18. A. C. Keene, S. Waddell, Drosophila olfactory memory: single genes to complex neural circuits. *Nature Reviews Neuroscience* **8**, 341–354 (2007).

19. G. M. Schratt *et al.*, A brain-specific microRNA regulates dendritic spine development. *Nature* **439**, 283–289 (2006).

20. S. Banerjee, P. Neveu, K. S. Kosik, A coordinated local translational control point at the synapse involving relief from silencing and MOV10 degradation. *Neuron* **64**, 871–884 (2009).

21. K. Tsurudome *et al.*, The Drosophila miR-310 Cluster Negatively Regulates

Synaptic Strength at the Neuromuscular Junction. *Neuron* **68**, 879–893 (2010).

22. C. McCann *et al.*, The Ataxin-2 protein is required for microRNA function and synapse-specific long-term olfactory habituation. *Proceedings of the National Academy of Sciences* **108**, E655–E662 (2011).

23. S. Bicker *et al.*, The DEAH-box helicase DHX36 mediates dendritic localization of the neuronal precursor-microRNA-134. *Genes \& Development* **27**, 991–996 (2013).

24. M. D. Paraskevopoulou *et al.*, DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Research* **41**, W169–73 (2013).

25. E. Berezikov *et al.*, Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. *Genome research* **21**, 203–215 (2011).

26. B. G. Dias, K. J. Ressler, Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat Neurosci* (2013), doi:10.1038/nn.3594.

27. M. Mirkovic-Hösle, K. Förstemann, T. Preiss, Ed. Transposon Defense by Endo-siRNAs, piRNAs and Somatic pilRNAs in Drosophila: Contributions of Loqs-PD and R2D2. *PLoS ONE* **9**, e84994 (2014).

28. N. V. Rozhkov, M. Hammell, G. J. Hannon, Multiple roles for Piwi in silencing Drosophila transposons. *Genes \& Development* **27**, 400–412 (2013).

29. P. N. Perrat *et al.*, Transposition-driven genomic heterogeneity in the Drosophila brain. *Science* **340**, 91–95 (2013).

30. B. Czech *et al.*, An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**, 798–802 (2008).

31. K. Okamura *et al.*, The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**, 803–806 (2008).

32. S. E. St Pierre, L. Ponting, R. Stefancsik, P. McQuilton, FlyBase Consortium, FlyBase 102--advanced approaches to interrogating FlyBase. *Nucleic Acids Research* **42**, D780–8 (2014).

33. J. Wen *et al.*, Diversity of miRNAs, siRNAs, and piRNAs across 25 Drosophila cell lines. *Genome research* **24**, 1236–1250 (2014).

34. M. He *et al.*, Cell-Type-Based Analysis of MicroRNA Profiles in the Mouse

Brain. *Neuron* **73**, 35–48 (2012).

35. E. Berezikov *et al.*, Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. *Genome research* **21**, 203–215 (2011).

36. P.-H. Wu, M. Isaji, R. W. Carthew, Functionally Diverse MicroRNA Effector Complexes Are Regulated by Extracellular Signaling. *Mol. Cell* **52**, 113–123 (2013).

37. K. Wibrand *et al.*, Differential regulation of mature and precursor microRNA expression by NMDA and metabotropic glutamate receptor activation during LTP in the adult dentate gyrus in vivo. *Eur J Neurosci* **31**, 636–645 (2010).

38. K. R. Nesler *et al.*, T. H. Gillingwater, Ed. The miRNA Pathway Controls Rapid Changes in Activity-Dependent Synaptic Structure at the Drosophila melanogaster Neuromuscular Junction. *PLoS ONE* **8**, e68385 (2013).

39. M. G. Thomas, M. L. Pascual, D. Maschi, L. Luchelli, G. L. Boccaccio, Synaptic control of local translation: the plot thickens with new characters. *Cell. Mol. Life Sci.* **71**, 2219–2239 (2014).

40. S. Murakami *et al.*, OptimizingDrosophilaolfactory learning with a semi-automated training device. *J Neurosci Methods* **188**, 195–204 (2010).

41. T. Ching, S. Huang, L. X. Garmire, Power analysis and sample size estimation for RNA-Seq differential expression. *RNA (New York, N.Y.)* **20**, 1684–1696 (2014).

42. W. Li *et al.*, MicroRNA-276a functions in ellipsoid body and mushroom body neurons for naive and conditioned olfactory avoidance in Drosophila. *The Journal of Neuroscience* **33**, 5821–5833 (2013).

43. A. N. A. Tayoun, C. Pikielny, P. J. Dolph, T. Zars, Ed. Roles of the Drosophila SK Channel (dSK) in Courtship Memory. *PLoS ONE* **7**, e34665 (2012).

44. H. Mi *et al.*, The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Research* **33**, D284–8 (2005).

45. S. L. Ameres *et al.*, Target RNA-directed trimming and tailing of small silencing RNAs. *Science* **328**, 1534–1539 (2010).

46. T. Katoh *et al.*, Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes \& Development* **23**, 433–438 (2009).

47. S. L. Fernandez-Valverde, R. J. Taft, J. S. Mattick, Dynamic isomiR regulation in

Drosophila development. *RNA* **16**, 1881–1888 (2010).

48. N. Cloonan *et al.*, MicroRNAs and their isomiRs function cooperatively to target common biological pathways. *Genome Biology* **12**, R126 (2011).

49. J. O. Westholm, E. Ladewig, K. Okamura, N. Robine, E. C. Lai, Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs. *RNA (New York, N.Y.)* **18**, 177–192 (2012).

50. B.-T. Juang *et al.*, Endogenous nuclear RNAi mediates behavioral adaptation to odor. *Cell* **154**, 1010–1022 (2013).

51. E. J. Lee *et al.*, Identification of piRNAs in the central nervous system. *RNA (New York, N.Y.)* **17**, 1090–1099 (2011).

52. N. R. Smalheiser, G. Lugli, J. Thimmapuram, E. H. Cook, J. Larson, Endogenous siRNAs and noncoding RNA-derived small RNAs are expressed in adult mouse hippocampus and are up-regulated in olfactory discrimination training. *RNA* **17**, 166–181 (2011).

53. Z. Yan *et al.*, Widespread expression of piRNA-like molecules in somatic tissues. *Nucleic Acids Research* **39**, 6596–6607 (2011).

54. modENCODE Consortium *et al.*, Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* **330**, 1787–1797 (2010).

55. K. Okamura, S. Balla, R. Martin, N. Liu, E. Lai, Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. *Nature Structural & Molecular Biology* **15**, 581–590 (2008).

56. Y. Kawamura *et al.*, Drosophila endogenous small RNAs bind to Argonaute|[thinsp]|2 in somatic cells. *Nature* **453**, 793–797 (2008).

57. B. Czech *et al.*, An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**, 798–802 (2008).

58. K. Okamura, N. Liu, E. C. Lai, Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. *Mol. Cell* **36**, 431–444 (2009).

59. K. Okamura *et al.*, The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**, 803–806 (2008).

60. M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406–3415 (2003).

61. C. J. Burke *et al.*, Layered reward signalling through octopamine and dopamine in Drosophila. *Nature* **492**, 433–437 (2012).

62. C.-L. Wu, M.-F. M. Shih, P.-T. Lee, A.-S. Chiang, An Octopamine-Mushroom Body Circuit Modulates the Formation of Anesthesia-Resistant Memory in Drosophila. *Current Biology* **23**, 2346–2354 (2013).

63. A. M. Burroughs *et al.*, A comprehensive survey of 3“ animal miRNA modification events and a possible role for 3” adenylation in modulating miRNA targeting effectiveness. *Genome research* **20**, 1398–1410 (2010).

64. S. L. Ameres, J.-H. Hung, J. Xu, Z. Weng, P. D. Zamore, Target RNA-directed tailing and trimming purifies the sorting of endo-siRNAs between the two Drosophila Argonaute proteins. *RNA (New York, N.Y.)* **17**, 54–63 (2011).

65. X. Li, I. M. Overton, R. A. Baines, L. P. Keegan, M. A. O'Connell, The ADAR RNA editing enzyme controls neuronal excitability in Drosophila melanogaster. *Nucleic Acids Research* **42**, 1139–1151 (2014).

66. J. Jepson, R. Reenan, Adenosine-to-Inosine Genetic Recoding Is Required in the Adult Stage Nervous System for Coordinated Behavior in Drosophila. *J. Biol. Chem.* **284**, 31391–31400 (2009).

67. G. Chawla, N. S. Sokol, ADAR mediates differential expression of polycistronic microRNAs. *Nucleic Acids Research* (2014), doi:10.1093/nar/gku145.

68. J. Yan *et al.*, The multiple-wing-hairs gene encodes a novel GBD-FH3 domain-containing protein that functions both prior to and after wing hair initiation. *Genetics* **180**, 219–228 (2008).

69. W. J. Gault, P. Olguin, U. Weber, M. Mlodzik, Drosophila CK1-γ, gilgamesh, controls PCP-mediated morphogenesis through regulation of vesicle trafficking. *The Journal of Cell Biology* **196**, 605–621 (2012).

70. S. Takayama *et al.*, Genome methylation in D. melanogaster is found at specific short motifs and is independent of DNMT2 activity. *Genome research* **24**, 821–830 (2014).

71. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25 (2009).

72. A. Kozomara, S. Griffiths-Jones, miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research* **39**, D152–7 (2011).

73. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

74. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

75. W. J. Kent *et al.*, The human genome browser at UCSC. *Genome research* **12**, 996–1006 (2002).

76. M. Robinson, D. McCarthy, G. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

Chapter III

Beta-Site APP-Cleaving Enzyme Is Required For Long Term Memory In *Drosophila*

Summary

Alzheimer's disease (AD) and related dementias are among the most widespread age related neurodegenerative diseases. Brains of AD patients exhibit tauopathy, accumulation of senile plaques, and cell death, resulting in progressive memory impairment and cognitive defects. Senile plaques are largely composed of a peptide termed amyloid-beta (Aβ), which is derived from cleavage of the Beta-Amyloid Precursor Protein (APP) by membrane-bound aspartic proteinases termed β-secretases. Such cleavage and resultant Aβ production were long thought to be essentially toxic, and to be the central disease mechanism underlying AD. However, recent work demonstrates that AD symptoms such as dementia are not always coupled to the presence of senile plaques in humans, and that Aβ is required for LTP in the mouse hippocampus. *Drosophila* express homologues of all of the genes thought to be central in AD progression. Using aversive olfactory classical conditioning, I show that LTM formation upregulates the Drosophila β-secretase dBACE, resulting in increased proteolytic processing of the APP homologue APP-Like (APPL). Further, using inducible RNAi knockdown of dBACE in the brain, we demonstrate that dBACE is required for LTM, but not for learning, or for STM.

Introduction

Memories are encoded in the synaptic connections of the brain's neural circuitry. An organism's ability to retain learned information over long periods of time, while remaining capable of acquiring new memories, requires a finely tuned balance of mechanisms that build, destroy, strengthen and weaken synapses. Misregulation of these processes can have profound effects on health, and can alter memory, cognition, and behavior. AD and related dementias are examples of how defects in the molecular mechanisms governing synaptic plasticity can cause profound changes in mental health. Common features of AD brains are tauopathy and the accumulation of the APP cleavage product Aβ into senile plaques. Tauopathy is a feature of many neurodegenerative diseases, while senile plaques are closely associated with Down syndrome and AD, both of which affect memory and cognition. As such, understanding the function of APP and its processing have been major areas of research into AD. Previous studies reveal that APP influences synaptic plasticity in complex ways, and that APP expression and processing must be exquisitely controlled for proper synaptic connectivity and memory.

The APP pathway is evolutionarily conserved from insects to humans. However, rodent Aβ does not form amyloid deposits. As such, researchers are forced to use exogenous human APP or Aβ in studies of AD in mice and rats. Recently, interest in *Drosophila* as a model organism in which to study the APP pathway and AD has grown, following the discovery that proteolytic processing of

the fly homologue of APP (APPL) yields a peptide functionally similar to Aβ (dAβ).(*1*) The *Drosophila* APPL pathway appears to function similarly to the human APP pathway, and many of the components from both species are mechanistically compatible. Expression of human APP rescues memory defects exhibited by *Drosophila* mutants of the APP homologue APPL.(*2*) The APPL processing machinery present in *Drosophila* cleaves human APP much as would occur in its endogenous setting. Overexpression of human Aβ in Drosophila leads to formation of aggregates resembling senile plaques, neurodegeneration, and memory defects. Co-overexpression of APPL and dBACE also leads to production of aggregates resembling senile plaques, despite the fact that APPL does not contain a sequence similar to Aβ. This demonstrates that the APP pathway is functionally conserved, even though some amino acid sequences of its components are not.(*1*) These findings and others show that *Drosophila* can serve as a model organism for understanding APP family protein metabolism and its function in synaptic plasticity and memory (reviewed in (*3*, *4*)).

APPL undergoes proteolytic processing along two pathways, each initiated by a characteristic cleavage event. Processing of APPL initiated by dBACE is considered amyloidogenic, as this pathway results in dAβ production and formation of deposits resembling senile plaques. APPL cleavage by dBACE liberates a soluble extrcellular N-terminal fragment, and a membrane bound C-terminal fragment (βCTF). Subsequent cleavage of βCTF within its transmembrane domain by the γ-secretase presenillin (Psn) produces dAβ. Due to the high degree of sequence

conservation within the CTF and functional conservation of the pathway, it is presumed that cleavage of βCTF by Psn also releases the APPL intracellular C-terminal domain (AICD). Drosophila AICD contains a highly conserved G_o-interacting domain, and the perfectly conserved YENPTY internalization sequence.(*5*) The high degree of sequence similarity between *Drosophila* and human AICD supports the assumption that they are functionally similar as well. In humans, AICD can translocate to the nucleus and induce transcriptional changes.(*3*, *4*) Though AICD production has been demonstrated in *Drosophila*, the resultant downstream effects remain uncharacterized.(*6*) Cleavage of APPL by the α-secretase Kuzbanian (Kuz) also liberates the N-terminal extracellular domain, and yields a C-terminal fragment (αCTF) that appears unable to generate dAβ or aggregates.(*1*) The APPL processing pathway initiated by Kuz is therefore termed nonamyloidogenic. In humans, APP processing is overwhelmingly initiated by α-secretase. However, AICD production is largely unaffected by α-secretase inhibitors. Furthermore, β-secretase initiated processing leads to greater nuclear accumulation of AICD.(*7-9*) These differences in AICD derived from α-γ-processing or β-γ-processing likely result from spatial segregation of α-secretase and β-secretase activity primarily to the cell surface and recycling endosomes, respectively. AICD is rapidly degraded in the cytosol. However, endosomes are transported along microtubules to the perinuclear area. AICD liberated by γ-secretase cleavage in this region is therefore less likely to be degraded before entering the nucleus than is AICD produced at the cell surface. This model is supported by the finding that manipulations driving endosome localization to the cellular periphery reduce AICD

nuclear accumulation.(*7*) Localization of β-secretase to endosomes also provides a mechanistic link between neural activity and AICD production. Action potentials trigger fusion of neurotransmitter containing vesicles with the plasma membrane, and subsequent clathrin dependent recycling endocytosis. This process leads to colocalization of APP previously present at the plasma membrane and β-secretase already residing within endosomes.(*10*) Thus, neural activity drives β-γ-secretase processing of APP. The functional conservation of the APP pathway in *Drosophila* suggests that β-γ-processing of APPL may also be stimulated by neural activity. Activity induced colocalization of dBACE and APPL could shift processing of APPL toward the amyloidogenic pathway, but this shift would be limited to the duration of increased synaptic vesicle release and recycling, were no other modulation of the pathway to occur. APPL levels and the relative abundance of its various metabolites can profoundly alter synaptic number and structure.(*5*) How APPL levels and the balance of the amyloidogenic and nonamyloidogenic pathways are altered or maintained remain poorly understood, and largely unstudied in *Drosophila*.

Here, I present results from our study of APPL processing in the course of LTM formation. This work was initiated after sRNA sequencing from head lysates of flies that had been subjected to olfactory aversive classical conditioning revealed a class of genes with an unusual sRNA signature. Genes in this class have a high degree of read coverage on the coding strand of exons. We therefore conclude that this signal is derived from mature transcripts, and designate this set as high exon coverage transcripts (HECTs). sRNA reads mapping to HECTs do not have features

of known sRNA classes such as esiRNAs, piRNAs, or microRNAs. A subset of HECTs display statistically significant changes in the HECT signal following conditioning, hinting at memory relevant regulation. Amongst regulated HECTs, dBACE has the largest change. I show that the increase in the dBACE signal is linked to increased dBACE expression and APPL processing. dBACE expression increases rapidly following LTM training, and remains elevated 24 hours later. This study is, to my knowledge, the first demonstration of activity dependent regulation of dBACE expression. I examine regulated HECTs for common sequence features, and show that intronless genes are overrepresented amongst upregulated HECTs, and underrepresented amongst downregulated HECTs. I show that most regulated HECTs carry a sequence element known to facilitate nuclear export of intronless transcripts via a pathway that involves components of the splicing machinery. I also present work from our lab demonstrating that dBACE expression is required for LTM, but not for short term memory (STM) or for learning. As far as we are aware, this is the first report of a requirement for dBACE in memory formation.

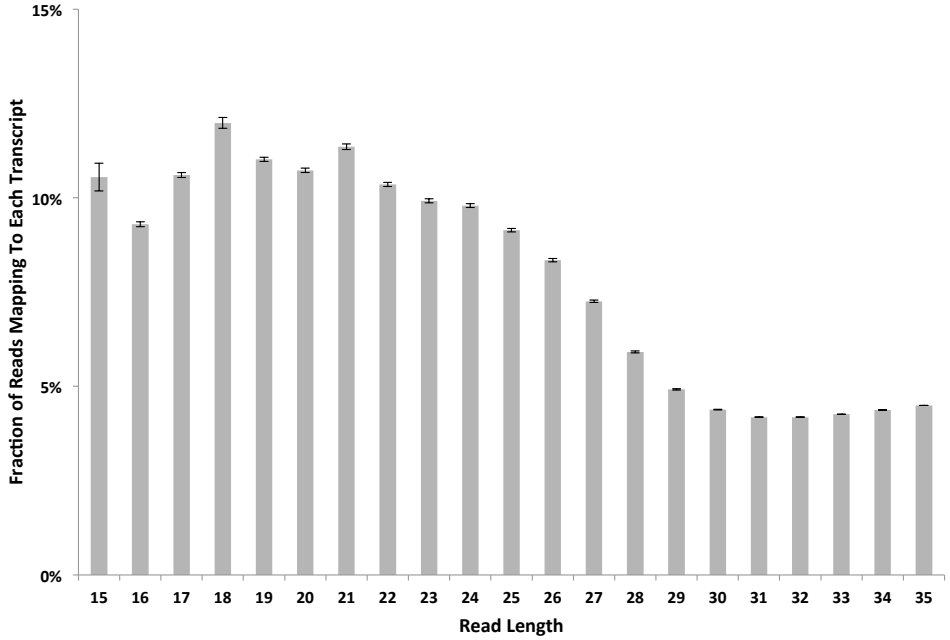
Results

sRNAs are produced from highly expressed transcripts

In chapter II, I describe experiments directed at characterizing changes in sRNA expression during LTM formation. In this effort, we conducted massively parallel sequencing of 15-35nt sRNAs extracted from heads of Drosophila subjected to aversive olfactory classical conditioning, or to the component stimuli (odor or electric shock) used in such training. While the primary goal was to profile microRNAs, esiRNAs and piRNAs, we also hoped to identify novel sRNA signals in our data. We aligned our sequencing reads to the *Drosophila* genome such that no mismatches were allowed, and all perfect alignments were retained. We then visualized these alignments using the UCSC genome browser.(*11*) Superficial visual inspection of our data using the UCSC browser revealed a set of genes exhibiting near complete read coverage across exons on the coding strand, and a near complete absence of reads mapping to introns. Further, a subset of these high exon coverage genes exhibits a large increase in reads vs. control in the LTM and/or shock only conditions. The near exclusive mapping of reads to exons of these genes suggests that these reads are derived from spliced mRNAs. We therefore remapped our reads to all spliced transcripts annotated in FlyBase. This step ensured that reads spanning exon junctions would be captured, and allowed us to more easily compare transcripts that include overlapping sets of exons. The hallmark features of high exon coverage transcripts (HECTs) are near complete coverage on the

coding strand, and a high sense:antisense read count ratio. Accordingly, we established >=95% read coverage, and a 10:1 sense:antisense read count ratio as parameters for designating transcripts as HECTs. We noted that some HECTs exhibit few reads at all in some experimental conditions. We therefore combined the reads from all of our sequencing libraries for the purposes of classifying transcripts as HECTs, so as to ensure that those HECTs displaying a dramatic difference in read counts between conditions would also be included in our subsequent analysis. These steps yielded 780 HECTs representing 465 genes (Figure S3.1). We next sought to determine if HECT reads had characteristics of known sRNA classes. Defined size ranges are a feature of known sRNAs. Accordingly, we computed read length distributions for each HECT. HECT reads have a much broader length distribution than miRNAs, which are 21~24nt, esiRNAs, which are almost exclusively 21-22nt, or piRNAs, which are largely 24-29nt (Fig 3.1).

Figure 3.1. Mean read length distribution for HECTs



Sequencing reads from all libraries were aligned to all FlyBase transcripts such that no mismatches were allowed, and all perfect alignments were retained. Those transcripts with >= 95% coverage on the coding strand and >= 10:1 sense:antisense read count ratio were designated as HECTs. Read length distributions were then computed for each HECT. HECT reads have a broad length distribution, unlike known sRNA classes. Bars represent mean read length distribution for HECTs. Error bars represent SEM.

This analysis indicates that HECT reads do not correspond to known sRNA classes. We note that many HECT genes are ubiquitously expressed, including Gapdh1, Gapdh2, Act5C, several ribosomal proteins, ATP synthases, and mitochondrial cytochrome c oxidase subunits. To test if ubiquitously expressed genes are in fact overrepresented in the set of HECT genes, we conducted gene ontology (GO) analysis using PANTHER.(*12*) This analysis shows that ribosomal proteins, Krebs cycle components, and cytoskeleton genes are overrepresented in the HECT gene set, while transcription factors are underrepresented (Figure 3.2).

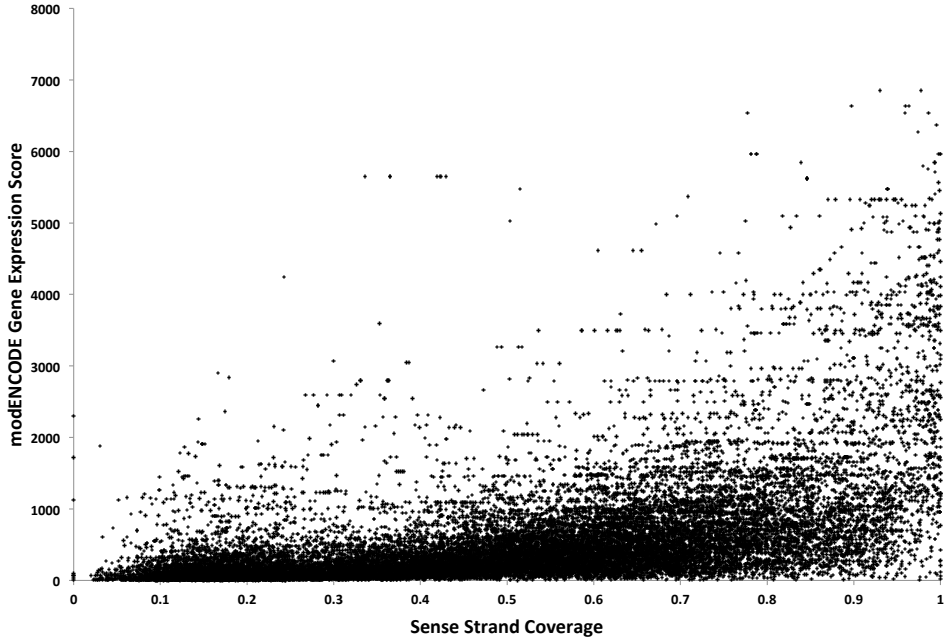
Figure 3.2. PANTHER gene ontology analysis of HECT genes

PANTHER Protein Class	Drosophila melanogaster	HECT Genes	Expected	over/under	P-value
ribosomal protein	164	61	5	+	2.08E-43
oxidoreductase	619	63	18.86	+	1.51E-14
dehydrogenase	252	36	7.68	+	7.77E-12
ATP synthase	44	13	1.34	+	3.23E-07
nucleic acid binding	1520	87	46.3	+	1.09E-06
transcription factor	740	3	22.54	-	3.95E-05
actin family cytoskeletal protein	169	20	5.15	+	7.48E-05

HECT gene GO annotations were obtained and tested for over/underrpresentaion using the PANTHER database. Ribosomal proteins, Krebs cycle components, and cytoskeleton genes are significantly overrepresented in HECT gene protein class GO annotations, while transcription factors are significantly underrepresented. P-values include Bonferroni correction for multiple testing.

This lead to the speculation that HECT reads might represent turnover products of highly expressed transcripts. We obtained gene expression level scores for all transcripts from modENCODE, and compared these scores to transcript coverage on the sense strand (Fig 3.3).

Figure 3.3 Comparison of sense strand read coverage and modENCODE gene expression score



Comparison of sense strand coverage and modENCODE gene expression scores. Sense strand read coverage from our combined libraries was computed for each FlyBase transcript. Gene expression scores for each transcript were obtained from modENCODE. Spearman rank correlation analysis shows that sense strand coverage is correlated with modENCODE gene expression scores (ρ=. 0.59, p-value < 2.2e-16).

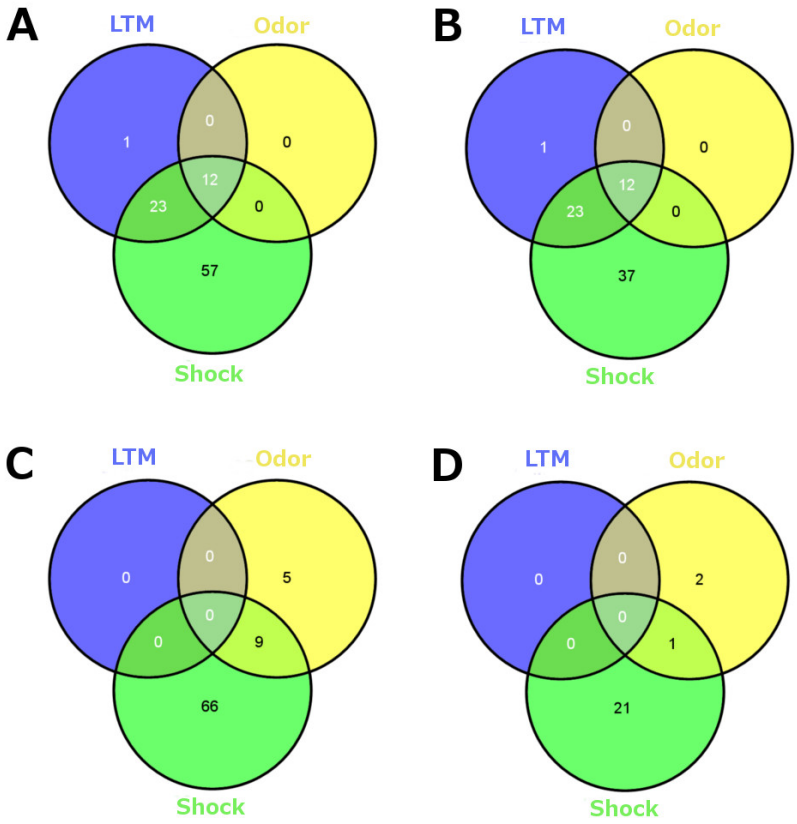
Spearman rank correlation analysis shows that sense strand coverage is correlated with modENCODE gene expression score (ρ=. 0.59, p-value < 2.2e-16). Though this analysis alone does not allow us to conclude that HECT reads represent turnover products of highly expressed transcripts, it does not contradict such a view.

All three experimental treatments induce significant changes in HECT read counts

If the HECT signature does represent turnover of highly expressed transcripts, then the subset of HECTs displaying large differences in read counts between our experimental conditions might reflect gene regulation in response to neural activity. To explore this possibility, we first tested HECTs for statistically significant differences in read counts from control using the edgeR software package.(*13*) 93 HECTs, representing 73 genes, display significantly increased read counts in the LTM, shock only or odor only conditions. 80 HECTs, representing 25 genes, display significantly decreased read counts in these experimental groups (Figure S3.2). While the shock only condition has the largest set, substantial overlap exists between experimental conditions in the sets of HECT genes with statistically significant differences in read counts from control (Fig 3.4). CG13324 is the lone HECT gene that has significantly increased reads only in the LTM condition. CG13324 reads also increase in the shock only condition, but this increase is not statistically significant. CG13324 is predicted to produce a 12.6kD polypeptide, but has no known function or conserved domains. No HECT genes have an increase in reads uniquely in the odor only condition. 37 HECT genes display a significant read

increase uniquely in the shock condition. All of these 37 HECT genes also have increased read counts in the LTM condition, though the increases in LTM are not statistically significant. 24 HECT genes have significantly decreased read counts in any condition, 21 of which are significant only in the shock only condition. No HECTs have significantly decreased read counts in the LTM condition.

Figure 3.4. Overlap between sets of HECT genes significantly regulated in each experimental condition.



HECT read counts were tested for significant differences from control using the edgeR software package. Venn diagrams representing overlap of HECTs or HECT genes found to differ significantly from control in each condition. A: Significantly upregulated HECTs. B: Significantly upregulated HECT genes. C: Significantly Downregulated HECTs. D: Significantly Downregulated HECT genes.

Proteases are enriched in the set of HECT genes with increased read counts in the LTM condition

We repeated our GO analysis on the set of HECT genes upregulated in any experimental condition. Genes involved in proteolysis are significantly overrepresented in this set (Fig 3.5).

Figure 3.5. Overrepresented molecular function annotations in the set of HECT genes significantly upregulated in any condition.

Molecular Function	Drosophila melanogaster	Up In Any Treatment Group	Expected	Over/Under	P-value
serine-type peptidase activity	301	15	1.59	+	8.10E-09
hydrolase activity	1629	29	8.61	+	1.67E-07
peptidase activity	591	17	3.12	+	1.48E-06
structural constituent of ribosome	157	8	0.83	+	3.05E-04
hydrolase activity, hydrolyzing O-glycosyl compounds	31	4	0.16	+	3.85E-03
amylase activity	15	3	0.08	+	1.19E-02
catalytic activity	3727	35	19.7	+	1.58E-02

Significantly overrepresented molecular function annotations for the set of HECT genes significantly upregulated in any condition were obtained using PANTHER. Genes annotated with molecular terms related to proteolysis are significantly overrepresented.

Nearly half (11 out of 23) HECT genes with significantly increased reads in both the LTM and shock only conditions, but not in the odor only condition, are annotated as serine type endopeptidases. Reads significantly increase for 12 HECT genes in all three experimental conditions vs control. 10 of these HECT genes have known or predicted functions, 8 of which are proteases. Ribosomal proteins are significantly overrepresented in the set of HECT genes with significantly increased reads only in the shock only condition. Many HECT genes with significantly decreased read counts in any condition are cytoskeletal or actin binding genes. Most of these genes have fewer reads in all conditions than those with significant increases.

Intronless genes are overrepresented in the set of HECT genes with significantly increased reads in the LTM condition.

The overrepresentation of proteases amongst HECT genes displaying increased reads in the LTM condition lead us to search for other common features of these genes. Superficial inspection using UCSC genome browser visualizations lead to the observation that many of these genes also lacked introns. To check whether this phenomenon is statistically significant, we compared the intron numbers of HECTs with significantly increased reads in each condition to the intron numbers of all HECTs, and of all Drosophila transcripts using hypergeometric distributions (Fig 3.6).

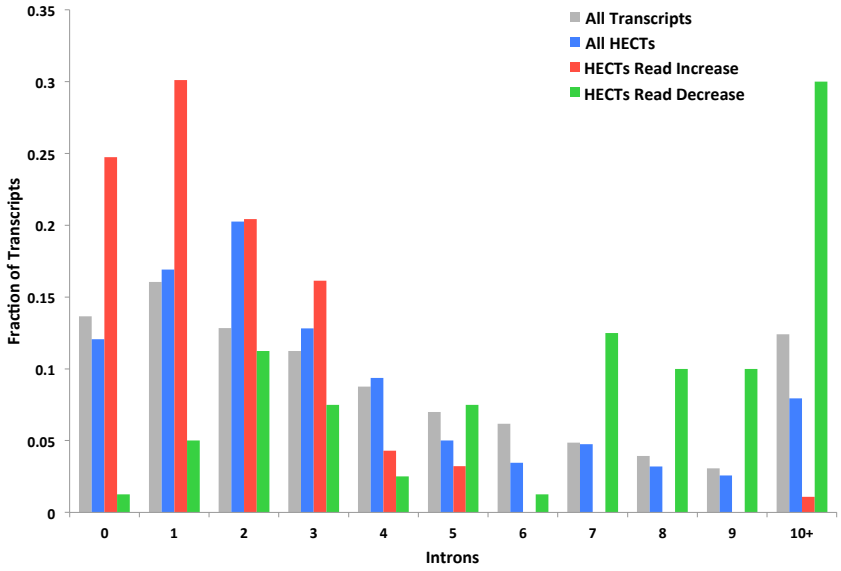
Figure 3.6. Comparison of intron numbers for HECTs and all transcripts.

A

	All Transcripts		All HECTs			Increased Reads			Decreased Reads		
Introns	Count	Fraction	Count	Fraction	pVal	Count	Fraction	pVal	Count	Fraction	pVal
0	4318	0.1366	94	0.1205	8.90E-02	23	0.2473	1.33E-02	1	0.0125	1.01E-03
1	5074	0.1606	132	0.1692	1.24E-01	28	0.3011	2.67E-03	4	0.0500	1.58E-02
2	4054	0.1283	158	0.2026	1.15E-08	19	0.2043	6.29E-02	9	0.1125	1.30E-01
3	3554	0.1125	100	0.1282	9.99E-02	15	0.1613	1.24E-01	6	0.0750	1.86E-01
4	2769	0.0876	73	0.0936	1.25E-01	4	0.0430	9.80E-02	2	0.0250	8.08E-02
5	2209	0.0699	39	0.0500	3.16E-02	3	0.0323	1.00E+00	6	0.0750	1.00E+00
6	1949	0.0617	27	0.0346	1.65E-03	0	0.0000	1.86E-02	1	0.0125	1.02E-01
7	1538	0.0487	37	0.0474	1.00E+00	0	0.0000	5.75E-02	10	0.1250	2.04E-02
8	1245	0.0394	25	0.0321	8.90E-02	0	0.0000	9.47E-02	8	0.1000	4.33E-02
9	970	0.0307	20	0.0256	6.33E-02	0	0.0000	5.48E-02	8	0.1000	1.55E-02
10+	3919	0.1240	62	0.0795	1.31E-04	1	0.0108	5.82E-04	24	0.3000	1.33E-04

Figure 3.6. (Continued)

B



C

	All Transcripts		All HECTs			LTM			Odor			Shock		
Introns	Count	Fraction	Count	Fraction	pVal	Count	Fraction	pVal	Count	Fraction	pVal	Count	Fraction	pVal
0	4318	0.1366	94	0.1205	8.90E-02	19	0.5429	5.12E-09	7	0.5833	1.44E-03	22	0.2391	3.38E-03
1	5074	0.1606	132	0.1692	1.24E-01	11	0.3143	1.31E-01	4	0.3333	8.26E-01	28	0.3043	3.04E-03
2	4054	0.1283	158	0.2026	1.15E-08	3	0.0857	2.57E-01	0	0	6.47E-01	19	0.2065	2.17E-01
3	3554	0.1125	100	0.1282	9.99E-02	0	0	7.31E-02	0	0	1.00E+00	15	0.163	2.84E-01
4	2769	0.0876	73	0.0936	1.25E-01	1	0.0286	6.74E-01	1	0.0833	1.00E+00	4	0.0435	1.95E-01
5	2209	0.0699	39	0.0500	3.16E-02	1	0.0286	9.23E-01	0	0	1.00E+00	3	0.0326	1.64E-01
6	1949	0.0617	27	0.0346	1.65E-03	0	0	1.00E+00	0	0	1.00E+00	0	0	2.22E-01
7	1538	0.0487	37	0.0474	1.00E+00	0	0	8.78E-01	0	0	1.00E+00	0	0	6.84E-02
8	1245	0.0394	25	0.0321	8.90E-02	0	0	6.23E-01	0	0	1.00E+00	0	0	2.06E-01
9	970	0.0307	20	0.0256	6.33E-02	0	0	3.95E-01	0	0	7.31E-01	0	0	2.36E-01
10+	3920	0.1241	62	0.0795	1.31E-04	0	0	2.14E-01	0	0	1.00E+00	1	0.0109	4.96E-03

D

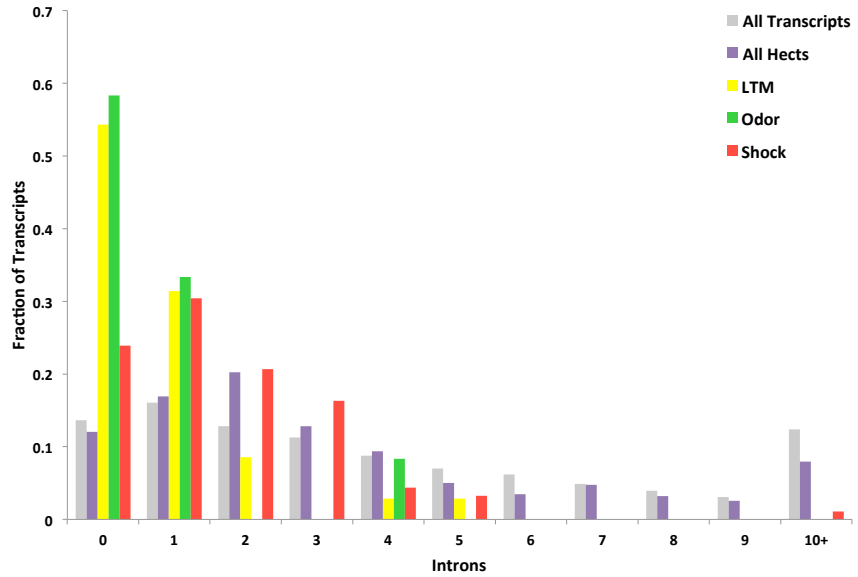
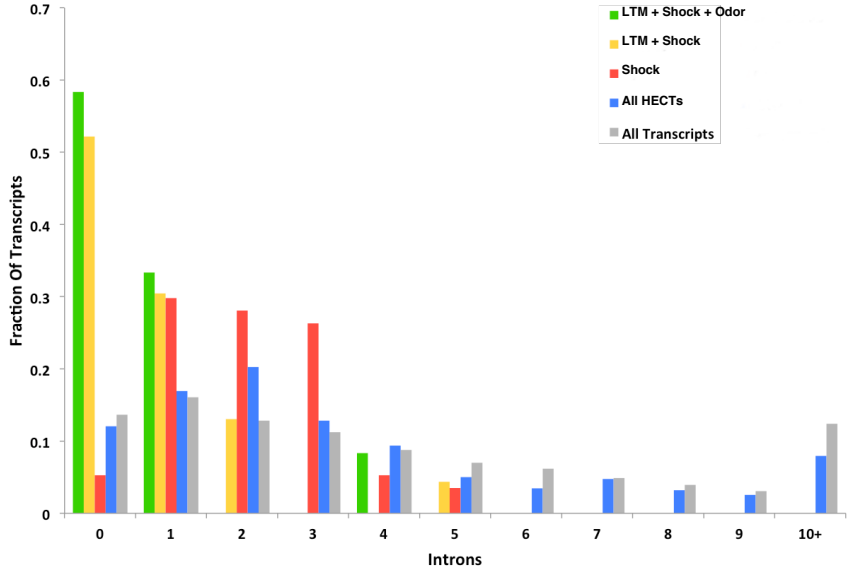


Figure 3.6. (Continued)

E



Intron counts for all *Drosophila* transcripts were obtained from FlyBase. Intron numbers of HECTs were compared using the hypergeometric distribution. P-values were corrected for multiple testing using the Benjamini-Hochberg method. A & B: While the intron count distribution for all HECTs largely resembles that of all transcripts, HECTs with 0 or 1 intron are over-represented amongst transcripts with increased reads following any treatment, while they are under-represented amongst transcripts with decreased reads following any treatment. C & D: Monoexonic transcripts are overrepresented amongst HECTs with increased read counts in the LTM, odor only, and shock only conditions. E: Grouping HECTs with significantly increased reads in more than one experimental condition reveals that the monoexonic overrepresentation in all conditions is due to the contribution of HECTs with a significant increase in the LTM condition.

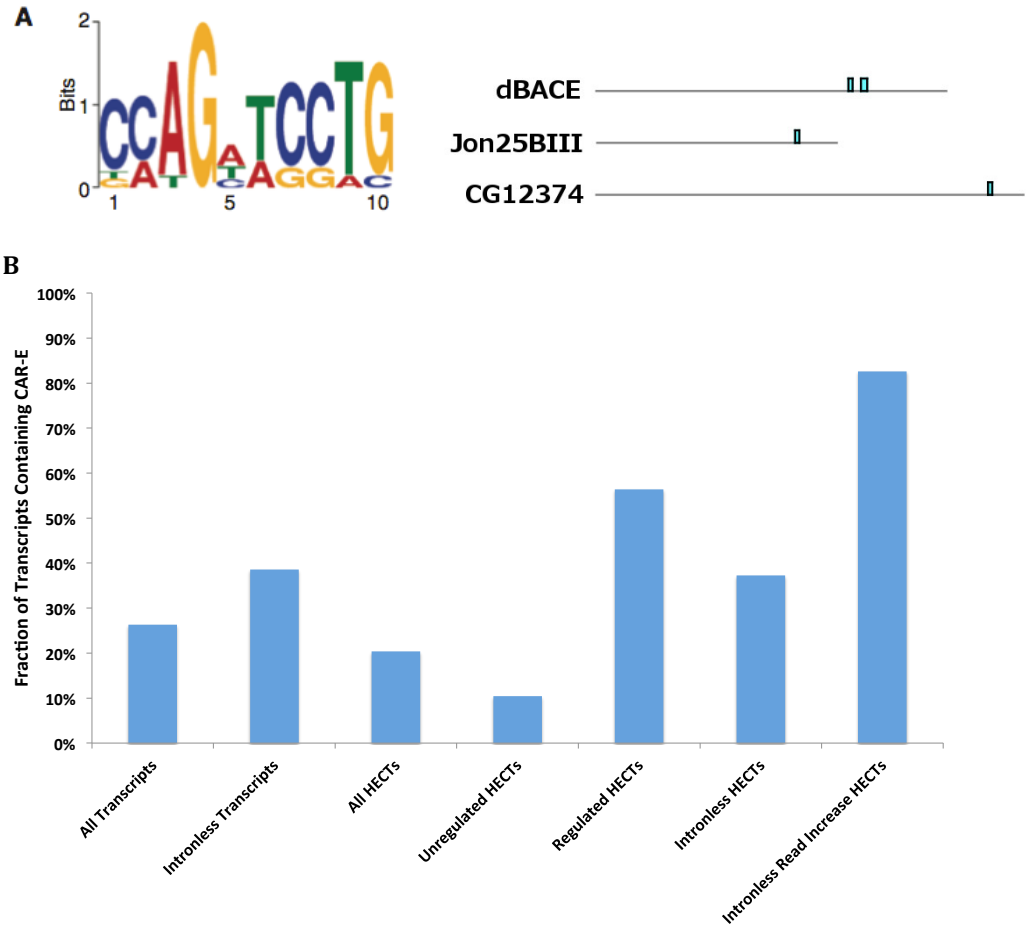
This analysis reveals that intronless transcripts are overrepresented amongst HECTs with significantly increased reads in any condition. However, examination of the overlap between conditions reveals that this overrepresentation is due to intronless HECTs with increased reads in the LTM condition (Fig.3.6 E).

Transcripts harboring a consensus sequence that facilitates nuclear export and expression of intronless mRNAs are overrepresented amongst regulated HECT genes.

We note that NMD, 5' cap formation, polyadenylation, and nuclear export are linked to splicing. Further, transcripts of naturally intronless genes, and cDNAs are less stable and are expressed at lower levels than transcripts that contain introns. This effect is tied to nuclear retention of transcripts that have not undergone splicing. In conditions of stress, expression of some intronless genes increases, while expression of genes with introns is reduced.(*14*) Surprisingly, the *Drosophila* splicing factor U2 small nuclear riboprotein auxiliary factor 50 (dU2AF^{50}) is required for nuclear export of many intronless transcripts.(*15*) In human cells, intronless genes and cDNAs harboring a consensus sequence element termed the cytoplasmic accumulation region element (CAR-E) are exported from the nucleus at substantially higher rates than those that do not contain the element. mRNAs bearing CAR-E tandem repeats assemble into RNPs containing the dU2AF^{50} homolog U2AF2.(*14*, *16*) The link between the CAR-E, U2AF2, and nuclear export suggests the possibility that such an element might also foster nuclear export of intronless mRNAs in *Drosophila*. If so, one might expect to find CAR-Es present in intronless HECTs. We used the FIMO and MAST algorithms to identify CAR-E occurrences in HECTs, and in all *Drosophila* transcripts.(*17*, *18*) This analysis revealed that CAR-E containing transcripts are slightly, though significantly underrepresented amongst all HECTs. However, further parsing of HECTs reveals CAR-E containing transcripts are significantly overrepresented amongst HECTs with significant changes in read counts in any of our treatment groups. Further, CAR-E containing transcripts are significantly underrepresented amongst HECTs that do not exhibit changes in read counts. Most dramatically, 19 out of 23 (82.61%, p =

7.53 E-8) intronless HECTs with increased reads in any treatment group contain at least one CAR-E (Fig3.7.).

Figure 3.7. CAR-E containing transcripts are overrepresented amongst regulated HECTs



	All Transcripts	Intronless Transcripts	All HECTs	Intronless HECTs	Unregulated HECTs	Regulated HECTs	Intronless Read Increase HECTs
Total	70127	2818	780	94	593	172	23
CAR-E	18458	1087	159	35	62	97	19
%	26.32%	38.57%	20.38%	37.23%	10.46%	56.40%	82.61%
Adj p-Val		1.93E-47	3.70E-05	5.79E-03	1.49E-21	2.15E-16	7.53E-08

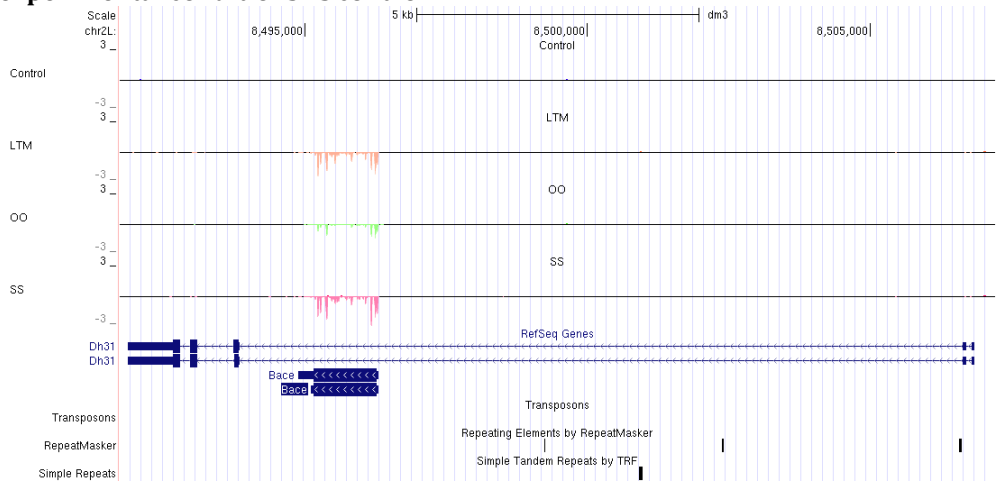
CAR-E containing transcripts are overrepresented amongst intronless HECTs with increased read counts following treatment. FIMO was used to identify CAR-E occurrences in transcripts. A: The left panel shows a sequence logo for CAR-E as reported in (*16*). The positions of CAR-E occurrences in the three HECTs with the largest read increases are shown in the right panel. B: Bar chart and table displaying the number of transcripts, and the number of transcripts containing at least one CAR-E in each category. P-values were computed using hypergeometric distributions, and were adjusted for multiple testing using the Holm-Bonferonni method.

Further investigation into whether CAR-Es, or U2AF^{50} are involved in the HECT signature, or in the increase in reads observed amongst intronless HECTs, was beyond the scope of this study. However, the strong overrepresentation of CAR-E containing transcripts amongst intronless HECTs that have increased reads following LTM training is intriguing.

dBACE mRNA is upregulated by LTM training, and by spaced sessions of the US alone

Our investigation of HECTs began with the observation that a large number of reads mapped to the dBACE gene in all experimental conditions, but not in control libraries. This signal was immediately apparent when visually examining our data using the UCSC genome browser, and inspired our search for other genes with this signature (Fig 3.8).

Figure 3.8. sRNA sequencing reads increase across the dBACE locus in all experimental conditions vs control

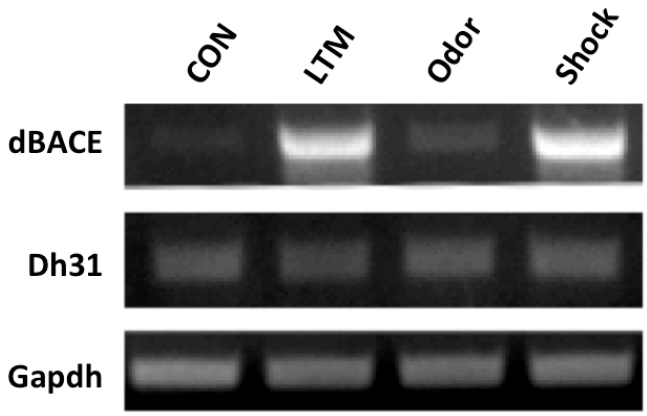


sRNA sequencing data visualized using the UCSC genome browser. Plots represent mean normalized reads mapping to each nucleotide position across the genome for each treatment group. Orange: LTM trained. Green: Odor only. Red: Shock Only. Few reads map within the Dh31 region in the control condition. The LTM, odor only, and shock only groups all exhibit significant read depth and coverage across dBACE gene on the coding strand.

Our subsequent differential expression analysis shows that dBACE is the HECT gene with the largest change in read counts vs control in all conditions (Figure S3.2). dBACE cleaves APPL, leading to production of dAβ. This pathway regulates neural morphology, synaptic plasticity, neurotoxicity, and behavior in insects, rodents, and humans.(*1*, *3*, *19*) We therefore sought to understand how the observed increases in dBACE reads relate to changes in dBACE expression. Semiquantitative RT-PCR conducted on the same total RNA samples used in preparing our sRNA sequencing libraries shows that dBACE mRNA is strongly upregulated in the LTM and shock only conditions, but not in the odor only condition. dBACE is an intronless gene residing within an intron of the Diuretic hormone 31 (Dh31) gene. Though few reads map to Dh31 exons in any condition, we also examined the possibility that

dBACE reads reflect Dh31 regulation or mRNA levels. Dh31 mRNA remained unchanged in all conditions (Fig. 3.9).

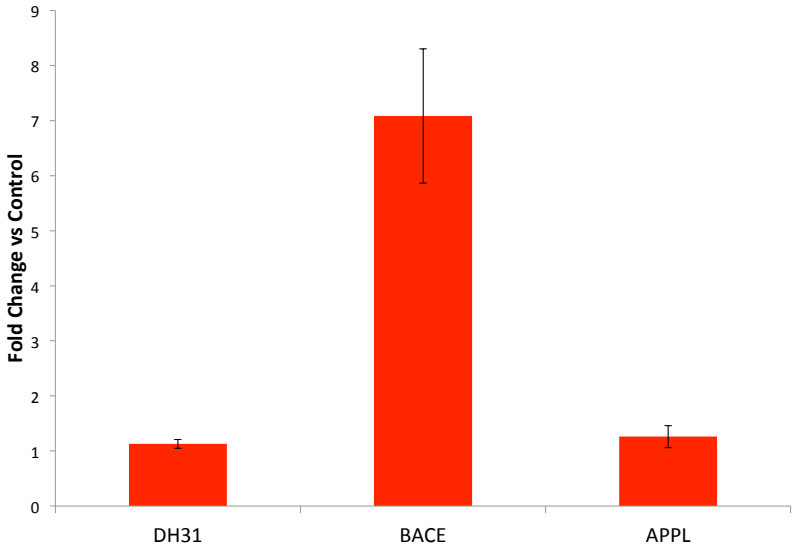
Figure 3.9. Semiquantitative RT-PCR reveals that dBACE mRNA is strongly upregulated in the LTM and shock only conditions



dBACE and Dh31 mRNA levels in the same total RNA samples used as input for sRNA sequencing library preparation were evaluated by semiquantitative RT-PCR. These samples were collected 2hrs after treatment. Poly-T oligonucleotides primed reverse transcription, and primers pairs spanning ~500bp exon sequences of dBACE, Dh31, or Gapdh were used for PCR amplification. dBACE mRNA is strongly upregulated in the LTM and shock only conditions, while it is only mildly induced in the odor only condition. Dh31 expression is not affected by any of our treatments.

Semiquantitative RT-PCR can give misleading results if the reaction is ended outside of the log phase of amplification. We therefore used quantitative real-time PCR analysis to better quantify the change in dBACE mRNA levels induced by LTM, and to ensure that Dh31 expression indeed remained unchanged. Further, dBACE expression might increase following LTM training to keep pace with increased APPL expression. Alternatively, elevated dBACE expression might reflect a shift to APPL processing along the amyloidogenic pathway. To investigate these possibilities, we examined APPL expression in these experiments as well. Again, we used the same total RNA samples from which our sRNA sequencing libraries were prepared (Fig. 3.10).

Figure 3.10. Real-Time PCR expression analysis following LTM training.



qRT- PCR shows that dBACE expression dramatically increases following LTM training, while APPL and Dh31 remain essentially unchanged. We used the same total RNA samples from which our sRNA sequencing libraries were prepared as input for these reactions. These samples were collected 2hrs after treatment. Poly-T oligonucleotides were used to prime reverse transcription. Primer pairs spanning ~500bp exon sequences were used to amplify Dh31, dBACE, APPL, or β-Tubulin. Triplicate technical repeats of each reaction were run on the same plate. This procedure was repeated for 3 independent matched sets of LTM trained and control RNA samples. Changes in expression were calculated using the ΔΔCt method, with β-Tubulin serving as the internal control. Error bars represent the standard error of the mean.

qRT-PCR analysis reveals that dBACE expression significantly increases (7.08 fold vs. control, p=2.48E-3) 2 hours after LTM training. Meanwhile, Dh31 and APPL mRNA levels remain essentially unchanged. These results suggest that at least at 2 hours after training, increased dBACE expression could reflect a change in APPL processing towards the amyloidogenic pathway, rather than a compensatory increase in dBACE expression to keep pace with elevated APPL expression. However, if Kuz expression were also increased, the relative preference for amyoidogenic vs non-amyloidogenic APPL processing could be maintained, even if total APPL cleavage rates increased. Furthermore, APPL expression, and the

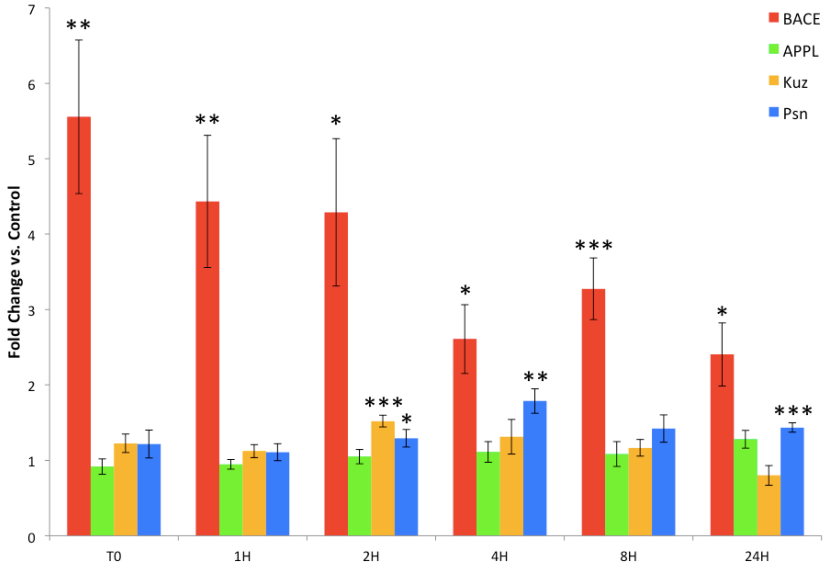
balance between APPL processing pathways may have important dynamics not apparent at the 2 hour post-training time point.

dBACE expression is rapidly upregulated following LTM training and remains elevated 24 hours post-training

The finding that dBACE is dramatically upregulated 2 hours after LTM training, or spaced sessions of US presentation, opened several lines of inquiry. First, how quickly after training is dBACE expression elevated, and how long does it stay elevated? Second, how do the other key components of the APPL pathway respond following training? The answers to these questions might provide insights into the roles of APPL and its metabolism in memory formation. We therefore conducted time course experiments, using qRT-PCR to simultaneously measure the expression of Kuz, dBACE, Psn, and APPL during the 24 hours following training. The fly handling and training of flies used in these experiments was nearly identical to that of our sRNA sequencing experiments. The only change in our procedures was that only LTM trained and control flies were prepared for the purposes of the time course experiments. We collected heads from 10 male and 10 female flies immediately after training, and again at 1, 2, 4, 8, and 24 hours after training. Total RNA was collected from LTM trained and control heads from each time point. We repeated these experiments 3 times, each time checking that a stable memory of the shock and odor pairing was present at 24 hours post-training. We performed each qRT-PCR measurement in triplicate on the same plate. Gapdh was used as an

internal control for all reactions. These experiments show that dBACE expression alone is strongly upregulated immediately after training, and decreases slowly over the course of 24 hours. Even with the decrease in expression that occurs after T0, dBACE remains elevated at the 24 hour time point (Fig. 3.11).

Figure 3.11. Time course measurement of APPL pathway component expression following LTM training.



qRT-PCR measurement of APPL pathway component expression shows that dBACE expression increases rapidly following LTM training, and remains elevated 24 hours later. Kuz expression increases mildly at the 2hr time point, but is otherwise relatively stable. Psn expression increases beginning 2hrs post training, and remains slightly, though significantly, elevated 24hrs post-training. APPL expression does not change significantly during the 24hrs following LTM training. Poly-T oligonucleotides were used to prime reverse transcription. Primer pairs spanning ~500bp exon sequences were used to amplify dBACE, APPL, Kuz, Psn, or gapdh. Triplicate technical repeats of each reaction were run on the same plate. This procedure was repeated for 3 independent matched sets of LTM trained and control RNA samples. Changes in expression were calculated using the ΔΔCt method, with gapdh serving as the internal control. Error bars represent the standard error of the mean. P-values were calculated using two-tailed student's t-test. (*: P<=0.05, **: P<=0.005, ***: P<=0.0005).

While dBACE is rapidly upregulated following training, expression of the other APPL pathway members we examined remains stable at the T0 and 1hr time points. At 2hrs post training, Kuz is mildly, though significantly upregulated (1.52 fold change,

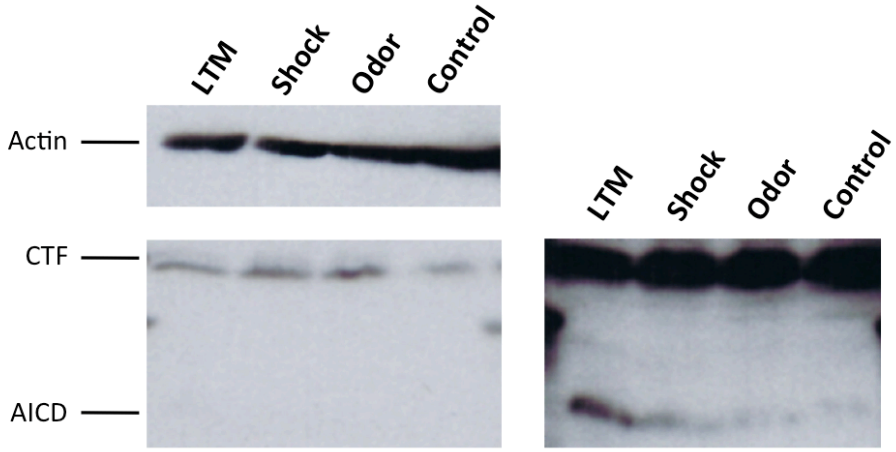
p<=.0005). At the 4hrs post-training time point, Kuz expression has declined somewhat, and is no longer significantly above control. From 4hrs to 24hrs post-training, Kuz expression continues to decline, but differences from control are not statistically significant. Psn expression also significantly elevated at the 2hrs post-training time point (1.29 fold change, p<=.05). Psn expression continues to rise up to ~1.8 fold (p<=0.005) above control at 4hrs post-training, and remains significantly elevated 24hrs after training (1.44 fold change, p<=0.0005). Though we observed a Psn expression level 1.42 fold higher than control at the 8hrs post-training time point, experimental variability meant that the difference was not statistically significant (p=0.60). Taken together, the time course qRT-PCR data indicate that expression of APPL pathway members is rapidly altered following LTM training, such that processing via the amyloidogenic pathway becomes more likely.

APPL processing is stimulated by LTM training and spaced sessions of US exposure.

We were not able to directly measure dBACE protein levels, as no antibody was available for such experiments. We speculated that increased dBACE expression might result in altered APPL processing. To test this idea, we examined APPL expression and processing in head lysates from flies subjected to LTM training, odor only, or shock only treatments and from control flies in the previously described manner using western blots. We used an antibody raised against a conserved C-terminal peptide within the AICD of *Manduca Sexta* APPL that is able to detect the C-termial end of *Drosophila* APPL to probe our western blots.(*1*, *20*)

These experiments show that all three experimental conditions mildly enhance a ~15kD band, a size corresponding to the CTFs resulting from the initial cleavage of APPL by either Kuz or dBACE. Long exposures reveal that LTM training, and to a lesser degree spaced sessions of US exposure, stimulate the appearance of an approximately 6kD band, a size corresponding to the predicted molecular weight of AICD (Fig 3.12).

Figure 3.12. LTM training and spaced sessions of shocks or odor enhance APPL processing



Western blots reveal that all three experimental treatments enhance APPL processing. Heads from flies in each condition were collected 2Hrs after treatment. Membranes were probed using α-msAPPL-AICD. Actin-5C was used as a loading control. In the left panel, a 4Hr exposure shows that samples from all three experimental conditions have slightly elevated levels of an ~15kD band. This size range corresponds to the C-terminal fragments of APPL resulting from cleavage by Kuz or dBACE. In the right panel, a 36Hr exposure of the same membrane reveals the presence of an additional ~6kD C-terminal APPL band. This size range matches the predicted molecular weight of the AICD. Both LTM training and spaced sessions of shocks increase the signal from this band, but the signal is much stronger in the LTM lane.

The 6 kD size range also corresponds to that predicted for dAβ. However, the antibody used in our western blots was raised against a peptide corresponding to the highly conserved YENPTY sequence present at the C-terminus of APPL. This

sequence resides within the AICD. Therefore, it is more likely that the 6 kD band corresponds to AICD than to dAβ.

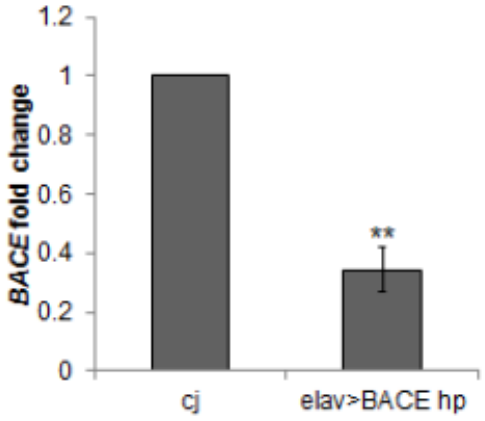
APPL and dBACE are required for aversive and appetitive LTM.

Having shown that training stimulates dBACE mRNA expression and APPL processing, we next sought to determine if memory formation is dependent on these activities. APPL knockout flies and flies overexpressing APPL or dBACE have been shown to have defects in perception of electric shock and locomotor defects that interfere with evaluation of memory using the T-maze test.(*1*) These defects are thought to be due to disturbances of neurodevelopment driven by such misexpression. We therefore sought to avoid such developmental defects by inducibly knocking down expression of APPL or dBACE just prior to training. To do so, we obtained Drosophila lines expressing RNA hairpins under UAS control directed against either dBACE (dBACE-hp) or APPL (APPL-hp). We then crossed the dBACE-hp or APPL-hp lines with enhancer trap flies expressing Gal4 driven by the panneuronal Elav enhancer, and a separate transgene in which a temperature sensitive variant of the yeast repressor Gal80 (Gal80ts) is expressed ubiquitously under the control of the tubulin promoter. In this set up, at normal temperatures Gal80ts blocks expression of the hairpins under UAS control. When flies are shifted to the restrictive temperature, Gal80ts is no longer able repress Gal4-UAS driven expression of the hairpins. The progeny of these crosses were subjected to aversive olfactory LTM training as previously described. Knockdown of either APPL or

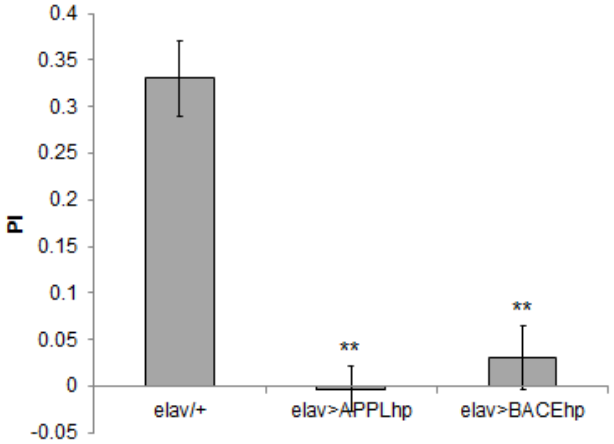
dBACE significantly reduces PIs when these flies were tested for olfactory LTM as previously described (Fig 3.13).

Figure 3.13. Knockdown of APPL or dBACE disrupt olfactory LTM

A



B



Flies harboring the panneuronal Elav-Gal4 driver, and the ubiquitously expressed tubulin-Gal80ts temperature sensitive repressor were used to conditionally express RNA hairpins directed against either APPL or dBACE. A)RT-PCR from heads of Elav-Gal4/Tub-Gal80ts/UAS-dBACE-hp flies housed at the permissive temperature, thus expressing dBACE-hp in all neurons, have significantly reduced dBACE mRNA levels. B)Flies in which panneuronal expression APPL-hp or dBACE-hp was induced for 30min immediately prior to training were subjected to aversive classical conditioning as previously described. Memory of the pairing of shock and odor was tested 24hrs later using a T-maze. APPL and dBACE hairpins both significantly reduced olfactory LTM PIs ($p<=0.005$).

To determine if the LTM defect caused by APPL-hp or dBACE-hp expression is generalizable to other training paradigms, we subjected the Elav-Gal4/tub-Gal80ts/UAS-APPL-hp and Elav-Gal4/tub-Gal80ts/UAS-dBACE-hp flies to appetitive olfactory conditioning. In this paradigm, flies learn to associate an odor with the availability of sucrose. Flies that have been deprived of food for 20hrs prior to training are exposed to a CS (odor) in the presence of a rewarding US (sucrose). These flies are also exposed to a control odor without US presentation. Memory is then evaluated in a T-maze as previously described. Using the appetitive conditioning method, a single 2 minute session produces protein synthesis dependent LTM.(*21*) Our experiments show that expression of APPL and dBACE are also required for appetitive olfactory LTM (Fig 3.14).

Figure 3.14. APPL or dBACE knockdown disrupts appetitive olfactory LTM
A

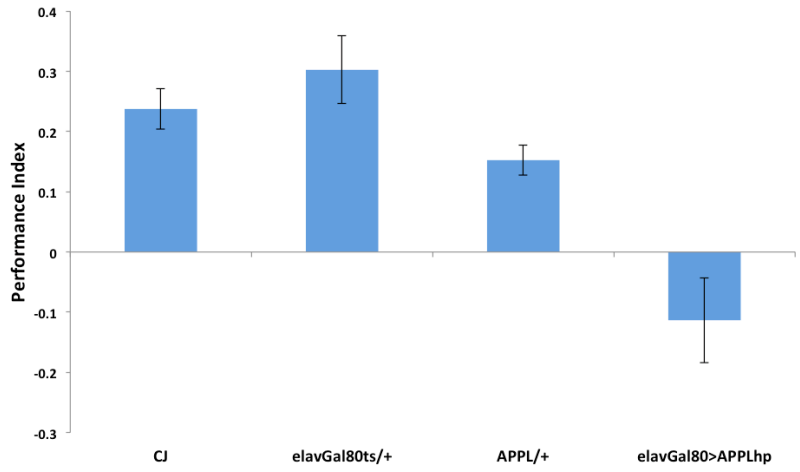
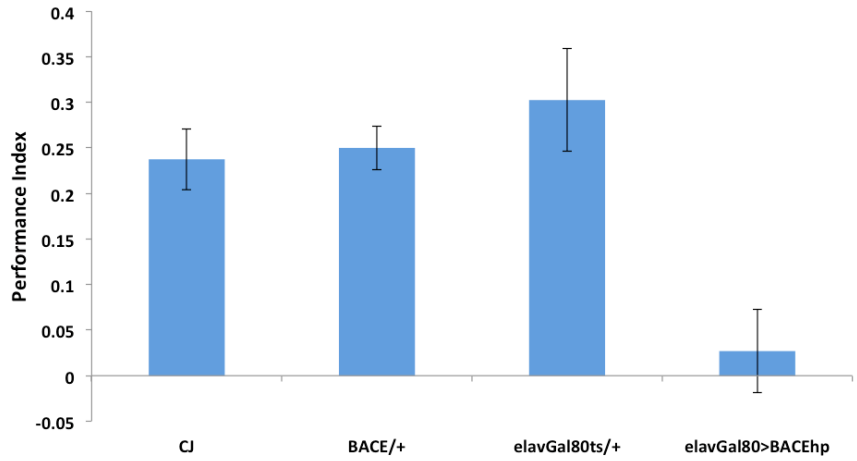


Figure 3.14. (Continued)

B



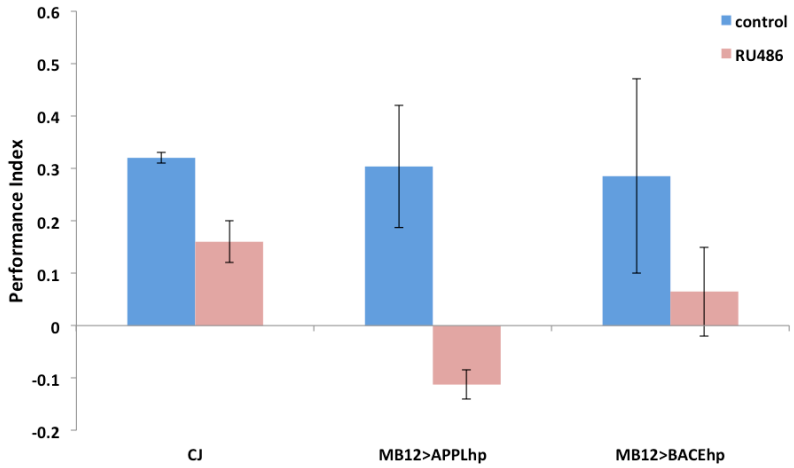
Flies in which panneuronal expression of RNA hairpins directed against either APPL or dBACE was induced just prior to training were subjected to appetitive olfactory classical conditioning. Memory of the pairing of sucrose and odor was tested 24hrs later using a T-maze. APPL (A) and dBACE (B) hairpins both significantly reduced appetitive olfactory LTM PIs.

Knockdown of APPL or dBACE in the adult MB disrupts LTM

The neural circuitry involved in *Drosophila* olfactory memory is well characterized. Lasting forms of both aversive and appetitive olfactory memory are encoded in the MB. Both involve dopaminergic signaling in the MB, but appetitive olfactory conditioning also involves octopaminergic input from APL neurons.(*22*, *23*) MB specific misexpression of some genes is sufficient to disrupt olfactory memory. Knockdown of APPL in the adult MB reduces LTM PIs in aversive olfactory classical conditioning.(*24*) To determine if appetitive olfactory LTM also requires APPL and dBACE expression in the MB, we tested flies expressing APPL-hp or dBACE-hp in the MB for LTM using our sucrose reward paradigm. For these experiments, we used the same gene switch system as was used in Goguel et al. to

induce hairpin expression within the MB beginning 48hrs prior to training. The gene switch system is a binary conditional expression method that drives transcription under the control of a UAS in response to administration of the drug RU486. A construct that encodes a chimeric protein harboring the DNA binding domain of the yeast Gal4 activator, the ligand binding domain of the mammalian progesterone receptor, and an activation domain from p65, is inserted into the genome. RU486 activates this chimeric protein, which is then able to drive transcription of UAS constructs already in widespread use. In the MB12 line, the gene switch construct is placed downstream of a 247bp Dmef2 enhancer element that is known to produce expression patterns that are restricted to the MB. The combination of this spatial restriction with temporal control through RU486 administration allows for precisely defined expression of UAS constructs.(*25*) MB12 driven expression of APPL-hp beginning 48hrs prior to appetitive conditioning produces impairment of olfactory LTM. dBACE-hp expression in the MB also reduces appetitive olfactory LTM, though not as strongly as APPL-hp (Fig. 3.15).

Figure 3.15. Expression of APPL-hp or dBACE-hp in the MB impairs LTM



Flies harboring UAS-APPL-hp or UAS-dBACE-hp constructs driven by the MB specific MB12 gene switch have lower performance indices when fed RU486 for 48hrs prior to appetitive olfactory conditioning than flies of the same genotype that were not fed RU486.

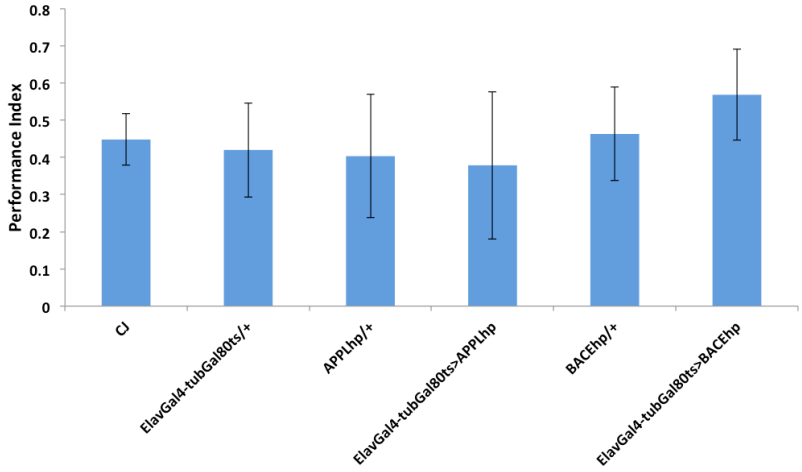
RU486 administration reduces PIs in control flies as well, but this effect alone does not account for the greater PI reductions in MB12>dBACE-hp or MB12>APPL-hp flies. Thus, our methods demonstrate that APPL expression is required in the MB for appetitive olfactory LTM, as well as for aversive olfactory LTM as reported by Goguel et al. Our data also indicate that dBACE-hp expression in the MB reduces appetitive olfactory LTM. However, the effect is smaller, and the negative impact of RU486 administration on PIs do not allow us to confidently conclude that the requirement for dBACE expression in olfactory LTM formation is confined to the MB.

APPL and dBACE are not required for STM

Our results showing that APPL is required for aversive and appetitive LTM are in agreement with a previous study in which APPL-hp was inducibly expressed

in the MB using the gene switch system.(*26*) Using this aproach, Goguel et al. showed that knockdown of APPL in adulthood reduces LTM, but leaves STM and perception of the US intact. Similarly, we find that adult expression of dBACE-hp in neurons also disrupts aversive and appetitive olfactory LTM. This result begs the question of whether neuronal dBACE-hp expression leaves STM in tact as is the case with APPL-hp. To test this, we subjected our Elav-Gal4/tub-Gal80ts/UAS-APPL-hp flies and Elav-Gal4/tub-Gal80ts/UAS-dBACE-hp flies to a single conditioning session pairing odor with shock after inducing hairpin expression. Such conditioning produces a memory of the pairing lasting on the order of hours in wild type flies. Two hours after conditioning, we tested for memory of the pairing using a T-maze as previously described. We find that neither APPL-hp or dBACE-hp significantly affects STM when expressed in neurons (Fig 3.16).

Figure 3.16. Panneuronal knock down of dBACE or APPL does not reduce STM



Flies were subjected to a single pairing of odor and shock and tested for memory of the pairing 2hrs after conditioning using a T-maze. STM scores for flies expressing APPL-hp or dBACE-hp in neurons were not significantly different from those of any of the control genotypes.

Discussion

During memory formation, synaptic efficacy and/or structure are altered. Formation of lasting memories requires that these alterations are durable. Transcription and translation are required for LTM, reflecting the need for altered gene expression programs to produce lasting changes in synaptic structure and function. After decades of study, novel aspects of gene regulation continue to be discovered. The widespread adoption of high-throughput nucleotide sequencing technology has lead to several such discoveries. Soon after their discovery, microRNAs, esiRNAs, and piRNAs were all shown to regulate important aspects of neuronal function and behavior.(*27-29*) With this in mind, we designed our sRNA study such that we would be able to profile these well studied sRNA classes, and to simultaneously extend our analysis beyond them. This approach yielded a wealth of unexpected sRNA sequences, many of which remain to be examined in depth. Of note, we identified a set of transcripts that produce copious reads mapping along the entirety of the sense strand of their exons. Though some of these sequences are present in previously published sRNA sequencing data sets, this study is to our knowledge the first to describe these HECT sRNAs as a class.(*30*) Further, we identify a subset of HECTs with read counts that change significantly during LTM formation, or following spaced sessions of exposure to odor alone or electric shock alone. Monointronic and intronless genes are overrepresented in the set of HECTs with increased reads following treatment and underrepresented amongst those with decreased reads. Furthermore, genes with GO annotations related to

proteolysis are also overrepresented amongst HECTs with increased read counts following training. These findings suggest the presence of a previously unrecognized regulatory mechanism driving increased expression of proteases following training.

While our primary goal was to profile microRNA expression during memory formation, the sensitivity and capacity of Illumina sequencing is several times greater than would be required to fully profile microRNA expression in a given sample. Rather than restricting our study to microRNAs and multiplexing our samples through barcoding, we chose to use the excess "sequencing space" to expand the size range of sRNAs studied. A size range of 21-29nt would have been sufficient to encompass esiRNAs, microRNAs, and piRNAs. However, other studies indicate that sRNAs outside of this size range exist in animal cells, and may be functional.(*31-33*) The 15-35nt size range we selected encompasses ~18nt transcription initiation RNAs (tiRNAs) at the low end, while remaining below the 36nt maximum size recommended in the Illumina sRNA library preparation kit v1.5 protocol. Reads with lengths outside the size ranges of microRNAs, esiRNAs, or piRNAs map to diverse loci across the genome. ncRNAs including rRNAs, snoRNAs, snRNAs, and tRNAs are major sources for such sRNA reads. sRNAs derived from ncRNAs have been reported previously, and aspects of their biogenesis and functions have been explored in a number of studies. (*34-39*) However, aside from snoRNA derived microRNAs, the functions of ncRNA derived sRNAs remain largely unknown in *Drosophila*. We also detected abundant tiRNAs at many transcription

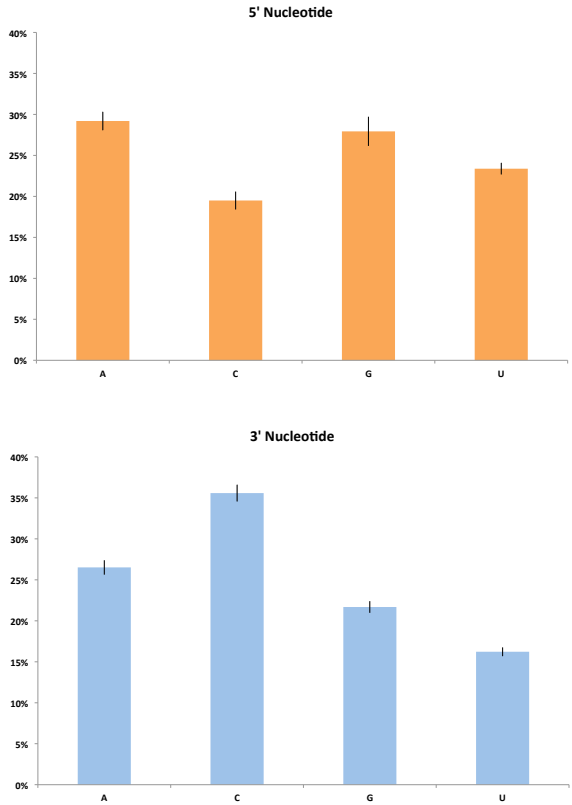
start sites (TSS). The number of unique tiRNA reads present at a TSS is correlated with expression levels of associated transcripts.(*32*, *40*) Though we note the presence of tiRNAs in our samples, analysis of these reads remains for future work. We also detected very short sRNA reads mapping near predicted insulator sites, and enhancers, but again, analysis of these reads is beyond the scope of this study. While we recognize the potential for novel insights in studying the classes of sRNAs just mentioned in our samples, we chose to focus on the abundant sRNA reads mapping to HECTs.

HECT sRNAs have a size distribution unlike those of other classes of sRNAs. MicroRNAs, esiRNAs, and piRNAs all have characteristic size ranges spanning 5nt or less. HECT sRNAs have a far broader size distribution, with 15nt-25nt reads constituting similar fractions of all sRNA reads. Our size selection precluded examination of RNAs shorter than 15nt, but shorter HECT sRNAs may also be present in animals. At the upper end of the HECT sRNA size range, reads longer than 26nt are less abundant, but 35nt reads are readily detectible. Longer HECT sRNAs may also occur, but again were outside of the range of our size selection, and thus do not appear in our data. The broad size range and transcript origin of HECT sRNAs suggests that they may be products of mRNA degradation. Previous studies in which parallel analysis of read ends (PARE) was used to explore the degradome of human cells suggest that this may be the case. The PARE method is used to determine the 5' end of cleaved or degraded mRNAs. PARE libraries are generated from polyadenylated mRNA fragments. The 5' PARE adapter is ligated using

chemistry that requires a 5' monophosphate on the poly-A tailed mRNA fragment. Such studies find monophosphorylated 5' ends mapping across transcripts. Bracken et al. show that higher PARE read counts are correlated with higher mRNA levels, but not with mRNA stability.(*41*) However some well expressed genes produce few if any PARE reads. As PARE relies on the presence of a 5' monophosphate, the absence of these well expressed genes in PARE libraries likely reflects degradation via mechanisms that do not leave a 5' monophosphate, and not necessarily greater mRNA stability.(*41*) Thus, existing evidence suggests that our analysis showing HECT read counts are correlated with modencode expression scores is indeed valid. Using PARE, read counts increase towards the 3' end of transcripts with multiple PARE reads, reflecting stalled processive 5'-3' degradation via the exonuclease Xrn1.(*41*, *42*) While we do not observe accumulation of reads at the 3' end of HECTs, this distinction could reflect either or both the underlying biology, and the differences between PARE and our sRNA sequencing methods. Reverse transcription during PARE library prep is primed by poly-T oligonucleotides. Thus, our methods capture products of endonuclease cleavage, and 3'-5' exonuclease activity that are not present in PARE libraries. Nonetheless, the patterns of sRNA read coverage across a given HECT is highly stereotyped amongst our samples. This may reflect aspects of the biogenesis of HECT sRNAs, but exploration of this idea will have to wait for future work. The preference for certain 5' or 3' sites for HECT sRNA reads of varying lengths does however support such a view. If HECT sRNAs are in fact produced via degradation of mRNAs, preferential mapping of 5' or 3' ends of reads to specific sites may be due to sequence or structural motifs involved in the

degradative process. We examined the possibility that mRNA secondary structure is related to HECT read coverage peaks using Mfold predictions for HECTs. This search was not comprehensive, but we were unable to identify any obvious relationship between structure predictions and read coverage peaks. We also looked for 5' and 3' nucleotide biases in HECT reads. We found that most HECT sRNAs have a 5' purine, and that cytidine is the least common 5' nucleotide. We also found that HECT sRNAs have a mild 3' bias for cytidine and against uridine. However, the significance of these biases is not immediately apparent. (Fig 3.17).

Figure 3.17: 3' and 5' nucleotide occurrence for HECT sRNAs



5' and 3' nucleotides were tallied for HECT sRNA reads in each library. These totals were divided by the total number of reads in the corresponding library. HECT sRNAs display a 5' preference for purines. Fewer sRNA reads have a 5' cytidine than any other nucleotide, while cytidine is the most common 3' nucleotide. Uridine is the least common 3' nucleotide.

Lastly, A+U rich elements (ARE) can affect the stability of their mRNA hosts. We conducted a search for AREs amongst 50 HECTs and did not find that transcripts containing AREs are overrepresented. Taken together, these results indicate that HECT sRNAs are not related to known classes of regulatory sRNAs, and are in fact degradation products of highly expressed mRNAs. However the nature of the processes that lead to HECT sRNA production remain to be elucidated in future work.

A subset of HECTs display statistically significant changes in read counts following training. Intriguingly, genes with GO annotations related to proteolysis are overrepresented in the set of HECTs with increased read counts. Further, HECT genes with increased read counts following treatment are largely intronless or have a single intron. Parsing this result further, we find that intronless genes are strongly overrepresented amongst HECTs with increased read counts in the LTM condition. Splicing and nuclear export are linked, and expression of intronless genes or transcripts of cDNAs is generally reduced in the absence of sequence motifs that allow for nuclear export via atypical mechanisms.(*14*) Previous work in human and Drosophila cells indicates that expression of intronless genes is regulated differently than genes whose transcripts undergo splicing. Conditions of cellular stress generally lead to translational arrest for most genes. However, in *Drosophila*, mRNAs for a set of intronless genes that are retained in the nucleus and thus poorly expressed under normal conditions, become cytoplasmicaly localized and escape translational silencing under conditions of stress.(*15*) In human cells, intronless

genes harboring the CAR-E consensus sequence in their coding regions are exported from the nucleus and expressed at far higher levels than those that do not contain a CAR-E.(*16*) We conducted a transcriptome wide search for the consensus CAR-E, and found that CAR-E containing transcripts are significantly overrepresented amongst regulated HECTs, and underrepresented amongst unregulated HECTs. Furthermore, we found that greater than 80% of intronless HECTs with increased read counts following training contain the consensus CAR-E. While these results strongly suggest that the CAR-E regulates expression of intronless HECTs following training, proof of such a model will depend on direct measurement of HECT gene expression and mutational analysis of the CAR-Es present in regulated HECTs.

The HECT gene with the largest read increase in all treatment groups was dBACE, a homologue of the human BACE genes. dBACE initiates amyloidogenic processing of APPL. In humans, misregulation of this pathway is thought to be the causal mechanism in AD progression. We find that increased dBACE reads reflect increased dBACE transcript levels. This increase corresponds to accumulation of APPL metabolites. To our knowledge, this study is the first to show that classical conditioning induces dBACE expression and APPL processing. Expression and proteolytic processing of APP family proteins influences neural morphology, synaptic structure, plasticity, and memory in complex ways. AD progression is typically associated with accumulation of senile plaques composed largely of Aβ. However, the view that β-secretase initiated processing of APP and resultant Aβ production is inherently toxic has been countered by recent studies in rodents and

humans (reviewed in (*43*)). Accumulating evidence indicates that APP expression and processing along the amyloidogenic and nonamyloidogenic pathways must all occur, and be exquisitely balanced for proper neural function and behavior. However, the difficulties of studies in humans and the absence of endogenous senile plaque forming Aβ in rodents has hampered efforts to understand the fundamental mechanisms by which APP controls synaptic plasticity, neural function, memory, and how these mechanisms are misregulated in AD. Recently, *Drosophila* has been shown to possess homologues of all of the key components of the APP pathway, and that APPL metabolism is functionally conserved. Importantly, β-γ-cleavage of APPL yields a peptide termed dAβ that forms aggregates similar to senile plaques and, though its sequence is not conserved, is functionally similar to Aβ.(*1*, *6*, *26*) These findings demonstrate that studies of APPL and its proteolytic processing can inform our understanding of APP processing and its role in AD. *Drosophila* behavioral genetics has proven instrumental in unraveling the biochemical underpinnings of many aspects of neural function, and indeed of memory itself. With these factors in mind, it seems likely that Drosophila will become an important model organism with which to study how processing of APP family proteins participates in memory formation. Yet, studies of APPL metabolism to date have largely relied on nonphysiologic conditions, and have focused on resulting pathologies. In this study, we examine the APPL pathway in wild type flies that have been subjected to aversive olfactory classical conditioning, or to exposure to the constituent odor (CS) or shock (US) alone. dBACE HECT sRNA reads increase significantly in the odor only, shock only, and LTM conditions. However, the increase is larger in the LTM

and shock only conditions than in the odor only condition, and semiquantitative RT-PCR shows that dBACE transcript levels are also induced much more strongly in the shock only and LTM conditions than in the odor only condition. Timecourse qRT-PCR shows that LTM training rapidly and strongly induces dBACE expression, but does not immediately alter expression of the other major APPL pathway components. At 24hrs post training, dBACE expression remains significantly elevated. By 2hrs post-training, the α-secretase Kuz is mildly induced, but this induction is fleeting, and Kuz expression is no longer significantly different from control at 4hrs post-training. Expression of the γ-secretase Psn is significantly elevated by 2hrs post-training, and remains elevated at 24hrs post-training. During the 24hrs following training, APPL expression remains largely unchanged. These qRT-PCR experiments suggest that APPL expression remains largely unchanged, but that its processing is increased and shifted towards the amyloidogenic β-γ-secretase pathway following training. If so, one would expect to observe elevated levels the β-CTF, Aβ, and AICD. However, previous studies have only reported the detection of β-CTF and AICD by Western blot through the use of transgenic or pharmacological interventions. Further, publications reporting the detection of dAβ and AICD by Western blot have utilized overexpression, epitope tags, or both in these experiments.(*1*, *6*) This reflects both the dearth of available reagents with which to study endogenous forms of these metabolites, and the low concentrations at which they are normally found. Accordingly, we were unable to clearly resolve distinct α-CTF and β-CTF bands in our Western blot experiments. We assume that this is due to the overwhelming preference for α-secretase cleavage of APPL and subsequent

obscurance of any β-CTF signal by the α-CTF in Western blot exposures long enough to detect the β-CTF at all. As described earlier in this chapter, AICD is highly unstable in *Drosophila* cells, and probably only accumulates when localized to the nucleus. Thus, the low molecular weight band we observe with very long western blot exposures following training may correspond to such nuclear accumulation. The ~6kD size could also correspond to the dAβ fragment, as both dAβ and AICD fall in this size range. However, the antibody we used was raised against a peptide corresponding to the conserved YENPTY motif present within AICD, and not in dAβ. While this fact gives us some confidence in asserting that this band is in fact AICD, additional experiments, perhaps using epitope tags, will be required to unambiguously identify it. In either case, the appearance of this band corresponds with the induction of dBACE, and likely reflects a shift to β-γ-secretase processing of APPL. Goguel et al. reported that APPL expression in the MB is required specifically for LTM, but not for learning, STM, or ARM.(*26*) Our data confirms this result, and expands upon it by showing that appetitive olfactory LTM also requires APPL expression in the MB. Further, we report reduced appetitive olfactory LTM in flies expressing dBACE-hp in the adult MB. However, the effect is not as strong as with APPL-hp, and the negative impact of RU486 on LTM prevents us from unambiguously concluding that dBACE-hp expression in the MB alone is sufficient to abrogate appetitive olfactory LTM. This result could be clarified through the use of other inducible gene expression systems that do not rely on RU486, such as Gal80ts. We also show that STM is retained in adult flies inducibly expressing APPL-hp or dBACE-hp in neurons. Importantly, ARM and LTM are distinguished by the

requirement for protein synthesis in LTM, but not in ARM. This leaves open the possibility that the ARM/LTM distinction reported by Goguel et al. may reflect a requirement for synthesis of dBACE protein and resultant elevated β-γ-cleavage of APPL during LTM formation. Alternatively, signaling downstream of APPL processing may drive production of other proteins required for LTM. Additional experiments will be required to make such an ARM/LTM distinction for flies expressing dBACE-hp, and to explore these possible explanations. Also, the relationship between dBACE expression, the appearance of the ~6kD APPL band following training, and LTM formation will need to be explored further. If this band is in fact AICD, and elevated dBACE activity leads to its accumulation, one would expect that this band is not produced following training in dBACE-hp flies. This experiment remains for future work. Deeper exploration of APPL cleavage following conditioning could benefit from a number of other systems with which to observe APPL cleavage and trafficking that are now available. In one instance, a construct in which Gal4 is fused to C99 (a C-terminal fragment of human APP), and a separate construct expressing a reporter under UAS control are introduced into the same fly. The UAS-reporter is only expressed when C99 is cleaved by Psn, and the AICD-Gal4 fusion enters the nucleus.(*44*) A similar construct fusing Gal4 to full length APPL could allow for visualization of APPL cleavage and resultant nuclear signaling during LTM formation. Others have used bimolecular fluorescence complementation to explore the protein-protein interactions involved in localizing AICD.(*45*) Future experiments using such tools could provide key insight into the function of dBACE driven APPL processing in response to classical conditioning.

The original goals of this study were centered around developing a better understanding of how known categories of short regulatory noncoding RNAs participate in memory formation, but our methods were intentionally designed with hopes of identifying novel mechanisms as well. These hopes motivated our decision to select for a wider RNA size range than was necessary to capture microRNAs, esiRNAs, and piRNAs, and to sequence at sufficient depth to capture meaningful information across that broad size range. Similarly, our analytical methods were not constrained to those already in use for a preconceived set of sRNAs. This approach lead to our identification of the unusual HECT sRNA signal, and ultimately to our exploration of dBACE's involvement in LTM formation. As such, this study is a clear demonstration of the value of genome wide survey experiments. Our data concerning the responses of microRNAs, piRNAs, and esiRNAs to classical conditioning are a valuable contribution to the knowledge base regarding sRNA function. However, the unforeseen identification of HECTs as a class, and the resulting series of experiments surrounding the activity of dBACE during LTM formation have direct bearing on a conserved mechanism known to be involved in AD. This study thereby supports the utility of *Drosophila* as a model organism with which to study AD. In the past, the relative ease and speed of fly work and the power of *Drosophila* genetics have been used to great effect in understanding the basic mechanisms of disease. Hopefully this will also be the case with AD. Further, the logical progression that led to our exploration of dBACE also yielded observations concerning the regulation of intronless genes that surely warrant

further experimentation. It may be that in casting a wide net, this study has opened a line of experimentation that will also contribute to our understanding of regulated nuclear export, and adds to the catalog of regulatory mechanisms involved in memory formation.

Materials and Methods

Drosophila rearing

Iso-CJ or CS-Quinn flies were reared on standard cornmeal medium at 25 °C under a 12hr light/dark cycle. 1-3 Day old adults were used in all experiments.

Aversive training and memory assay

Training was conducted using a semi automated conditioning apparatus as described in (*9*). Briefly, a single batch of flies was split into one of four groups: Control, odor only, shock only, or shock + odor (LTM), and housed in bottles containing no food for one hour prior to training. Flies were then loaded into vials that permit the flow of air from the apparatus and contain electrode grids. Odor and shock delivery are controlled by a computer. Following training, flies were returned to bottles containing no food, and allowed to rest for 2 hours before dissection. Memory formation in LTM flies was evaluated using a T-maze apparatus as described in (*8*). The performance index (PI) was calculated as the number of flies

avoiding the conditioned odor minus the number of flies avoiding a control odor divided by the total number of flies in the experiment.

Appetitive training and memory assay

48 hrs before training, flies were pre-starved by transferring them into bottles without food, but containing a kimwipe on which 1.5ml of 2% sucrose had been soaked. Flies were then starved for 24hrs prior to training by transferring them to a bottle containing only a dampened kimwipe. Pre-starvation and starvation occurred at 25° C, unless the temperature sensitive Gal80ts allele was being induced, in which case starvation occurred at 30° C. Training occurred in tubes containing either dampened filter paper (control) or filter paper containing 2M sucrose solution (LTM trained). First, all flies are placed into control tubes, and exposed to the CS- odor for 2min. Flies are then transferred into new tubes. Control flies are transferred to a new control tube, and LTM trained flies are transferred to a tube containing filter paper soaked in sucrose. Flies are then exposed to the CS+ odor for 2min. At this point, the handling of flies for LTM testing and STM testing will differ. For LTM, flies are transferred to regular food for 2 hrs and then move to a starvation bottle again for 18-20 hrs before testing. For STM, flies are transferred to an empty tube and connect to training apparatus and exposed to fresh air for 2 min. Memory formation was evaluated using a T-maze apparatus as described in (*8*). The performance index (PI) was calculated as the number of flies avoiding the conditioned odor minus the number of flies avoiding a control odor divided by the total number of flies in the experiment.

Tissue processing and total RNA extraction

For each sample, heads of 10 male and 10 female flies were dissected and immediately flash frozen and stored at -80 °C while awaiting results from memory testing. Only samples from matched batches of flies in which the LTM trained flies had a PI >=20 were used in subsequent steps. Heads were homogenized in Qiazol (Qiagen) using a motorized tissue homogenizer. Total RNA was extracted using the Qiagen miRNEasy micro kit, according to the manufacturer's protocol. RNA concentration was obtained using a Nanodrop and Bioanalyzer. RNA quality was determined by Bioanalyzer analysis.

sRNA sequencing

Prior to sequencing library preparation, total RNA used for sRNA sequencing was size selected by PAGE, as recommended in the Illumina sRNA sample prep kit (v1.5) protocol. Gel slices containing only 15-35nt RNAs were excised and used in downstream library preparation steps. sRNA sequencing libraries were prepared using the Illumina sRNA sample prep kit (v1.5) according to the manufacturer's protocol. Library quality and concentration were determined by Bioanalyzer analysis. sRNA sequencing was conducted by Harvard systems biology core facility staff, using an Illumina Genome Analyzer IIx.

Sequencing data analysis

Adapter sequences were removed using the FASTX toolkit. For microRNA analysis, sequencing reads were aligned to miRBase release 19 Drosophila hairpin sequences using Bowtie v0.12.7.(*45*, *46*) Custom software written in the R and AWK languages were used for canonical and isomiR analysis. For esiRNA and piRNA analysis, reads were aligned to the FlyBase Drosophila melanogaster genome (v5.48) using Bowtie v0.12.7.(*45*). SAMtools was used to convert alignments to BAM format.(*47*) BedTools v2.18.2 was used to count reads aligning within a given feature, and to merge reads into read-contigs.(*48*) The UCSC genome browser was used to visualize sequencing data.(*49*)

Differential expression analysis

The Bioconductor package edgeR v3.1.9 was used in its GLM mode for read count differential expression analysis. Read counts were normalized using the upper quartile method.(*50*) For isomiR analysis, the percentage of normalized reads with a given non-canonical feature was computed for each miRNA. Two tailed students t-tests were used to test for statistical significance of differences in these percentages. P-values were then adjusted for multiple testing using the Holm-Bonferroni method.

Gene ontology analysis

The PANTHER gene ontology (GO) database was used to assign GO terms and search for over/underrepresented terms.(*14*)

Genomic information

Genome wide annotations, including intron counts, were obtained from FlyBase (r5.48).

RNA secondary structure prediction

Mfold was used to predict RNA secondary structures.(*33*)

Western blot

For each sample, heads of 10 male and 10 female flies were dissected and immediately flash frozen and stored at -80 °C for 12hrs. Heads were homogenized in Laemmli buffer containing protease inhibitors (Bio-Rad) using a motorized tissue homogenizer. Samples were electrophoresed on 4-20% gradient Tris-HCl-acrylamide gels (Bio-Rad) and transferred to PVDF membranes (Bio-Rad).

Semiquantitative RT-PCR

First strand cDNA was reverse transcribed from 500ng input total RNA, using poly-T primers (Invitrogen) and the Superscript II kit (Invitrogen). Reverse transcription occurred at 42 °C for 1 hour. Oligonucleotide pairs spanning transposon sequences were used to prime 20 cycles of PCR amplification. Perfect Taq (5 prime) polymerase was used for PCR amplification.

qRT-PCR

First strand cDNA was reverse transcribed from 500ng input total RNA, using poly-T primers (Invitrogen) and the Superscript II kit (Invitrogen). Reverse transcription

occurred at 42 °C for 1 hour. Oligonucleotide pairs spanning ~500nt transcript sequences were used to prime amplification. iTaq™ universal SYBR® Green supermix (Bio-Rad) was used for real-time amplification. The reactions were incubated in a 96-well optical plate at 95°C for 5 min, followed by 40 cycles of 95°C for 15 s and 60°C for 30 s. The threshold cycle (Ct) values were determined using default threshold settings. The Ct is defined as the fractional cycle number at which the fluorescence passes the fixed threshold. For each RNA preparation, qRT-PCR reactions (including no template control), were performed in triplicate. The Ct values determined for each sample were normalized to the average Ct obtained for Gapdh, calculated from triplicate reactions. qRT-PCR was performed using an MJ Opticon (PTC-200 DNA Engine Cycler and CFD-3200 Opticon Detector) from Bio-Rad (formerly MJ research).

PCR Primers

Literature Cited

1. K. Carmine-Simmen *et al.*, Neurotoxic effects induced by the Drosophila amyloid-beta peptide suggest a conserved toxic function. *Neurobiology of Disease* **33**, 274–281 (2009).

2. L. Luo, T. Tully, K. White, Human amyloid precursor protein ameliorates behavioral deficit of flies deleted forapplgene. *Neuron* **9**, 595–605 (1992).

3. N. N. Nalivaeva, A. J. Turner, The amyloid precursor protein: A biochemical enigma in brain development, function and disease. *FEBS Letters* **587**, 2046–2054 (2013).

4. B. Poeck, R. Strauss, D. Kretzschmar, Analysis of amyloid precursor protein function in Drosophila melanogaster. *Exp Brain Res* **217**, 413–421 (2012).

5. L. Torroja, M. Packard, M. Gorczyca, K. White, V. Budnik, The Drosophila beta-amyloid precursor protein homolog promotes synapse differentiation at the neuromuscular junction. *The Journal of Neuroscience* **19**, 7793–7803 (1999).

6. C. Groth, W. G. Alvord, O. A. Quinones, M. E. Fortini, Pharmacological analysis of Drosophila melanogaster gamma-secretase with respect to differential proteolysis of Notch and APP. *Molecular Pharmacology* **77**, 567–574 (2010).

7. Z. V. Goodger *et al.*, Nuclear signaling by the APP intracellular domain occurs predominantly through the amyloidogenic processing pathway. *J. Cell. Sci.* **122**, 3703–3714 (2009).

8. B. B. Flammang *et al.*, Evidence that the amyloid-β protein precursor intracellular domain, AICD, derives from β-secretase-generated C-terminal fragment. *J Alzheimers Dis* **30**, 145–153 (2012).

9. N. D. Belyaev *et al.*, The transcriptionally active amyloid precursor protein (APP) intracellular domain is preferentially produced from the 695 isoform of APP in a {beta}-secretase-dependent pathway. *J. Biol. Chem.* **285**, 41443–41454 (2010).

10. U. Das *et al.*, Activity-Induced Convergence of APP and BACE-1 in Acidic Microdomains via an Endocytosis-Dependent Pathway. *Neuron* **79**, 447–460 (2013).

11. W. J. Kent *et al.*, The human genome browser at UCSC. *Genome research* **12**, 996–1006 (2002).

12. H. Mi *et al.*, The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Research* **33**, D284–8 (2005).

13. M. Robinson, D. McCarthy, G. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

14. R. Reed, E. Hurt, A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* **108**, 523–531 (2002).

15. M. Blanchette, E. Labourier, R. E. Green, S. E. Brenner, D. C. Rio, Genome-wide analysis reveals an unexpected function for the Drosophila splicing factor U2AF50 in the nuclear export of intronless mRNAs. *Molecular Cell* **14**, 775–786 (2004).

16. H. Lei, B. Zhai, S. Yin, S. Gygi, R. Reed, Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export. *Nucleic Acids Research* **41**, 2517–2525 (2013).

17. T. L. Bailey, M. Gribskov, Combining evidence using p-values: application to sequence homology searches. *Bioinformatics* **14**, 48–54 (1998).

18. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).

19. B. J. Bolkan, T. Triphan, D. Kretzschmar, β-secretase cleavage of the fly amyloid precursor protein is required for glial survival. *The Journal of Neuroscience* **32**, 16181–16192 (2012).

20. L. M. K. T. M. C. S. M. F. M. A. S. P. F. C. T L Swanson, THE INSECT HOMOLOGUE OF THE AMYLOID PRECURSOR PROTEIN INTERACTS WITH THE HETEROTRIMERIC G PROTEIN Goα IN AN IDENTIFIED POPULATION OF MIGRATORY NEURONS. *Developmental biology* **288**, 160–178 (2005).

21. M. J. Krashes, S. Waddell, Drosophila Appetitive Olfactory Conditioning. *Cold Spring Harb Protoc* **2011**, pdb.prot5609–pdb.prot5609 (2011).

22. C. J. Burke *et al.*, Layered reward signalling through octopamine and dopamine in Drosophila. *Nature* **492**, 433–437 (2012).

23. C.-L. Wu, M.-F. M. Shih, P.-T. Lee, A.-S. Chiang, An Octopamine-Mushroom Body Circuit Modulates the Formation of Anesthesia-Resistant Memory in Drosophila. *Current Biology* **23**, 2346–2354 (2013).

24. V. Goguel *et al.*, Drosophila amyloid precursor protein-like is required for long-term memory. *The Journal of Neuroscience* **31**, 1032–1037 (2011).

25. Z. Mao, G. Roman, L. Zong, R. L. Davis, Pharmacogenetic rescue in time and space of the rutabaga memory impairment by using Gene-Switch. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 198–203 (2004).

26. V. Goguel *et al.*, Drosophila Amyloid Precursor Protein-Like Is Required for Long-Term Memory. *The Journal of Neuroscience* **31**, 1032–1037 (2011).

27. P. Rajasethupathy *et al.*, A Role for Neuronal piRNAs in the Epigenetic Control of Memory-Related Synaptic Plasticity. *Cell* **149**, 693–707 (2012).

28. G. Siegel, R. Saba, G. Schratt, microRNAs in neurons: manifold regulatory roles at the synapse. *Curr. Opin. Genet. Dev.* **21**, 491–497 (2011).

29. B.-T. Juang *et al.*, Endogenous nuclear RNAi mediates behavioral adaptation to odor. *Cell* **154**, 1010–1022 (2013).

30. modENCODE Consortium *et al.*, Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* **330**, 1787–1797 (2010).

31. A. Sobala, G. Hutvagner, Small RNAs derived from the 5' end of tRNA can inhibit protein translation in human cells. *RNA Biol* **10**, 553–563 (2013).

32. R. Taft, P. Hawkins, J. Mattick, K. Morris, The relationship between transcription initiation RNAs and CCCTC-binding factor (CTCF) localization. *Epigenetics & Chromatin* **4**, 13 (2011).

33. P. Kapranov *et al.*, New class of gene-termini-associated human RNAs suggests a novel RNA copying mechanism. *Nature* **466**, 642–646 (2010).

34. M. Couvillion, G. Bounova, E. Purdom, T. Speed, K. Collins, A Tetrahymena Piwi Bound to Mature tRNA 32 Fragments Activates the Exonuclease Xrn2 for RNA Processing in the Nucleus. *Mol. Cell* **48**, 509–520 (2012).

35. Z. Li *et al.*, Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Research* **40**, 6787–6799 (2012).

36. T. Hanada *et al.*, CLP1 links tRNA metabolism to progressive motor-neuron loss. *Nature* **495**, 474–480 (2013).

37. S. Yamasaki, P. Ivanov, G.-F. Hu, P. Anderson, Angiogenin cleaves tRNA and promotes stress-induced translational repression. *The Journal of Cell Biology* **185**, 35–42 (2009).

38. P. Kumar, J. Anaya, S. B. Mudunuri, A. Dutta, Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biology* **12**, 78 (2014).

39. C. Ender *et al.*, A human snoRNA with microRNA-like functions. *Mol. Cell* **32**, 519–528 (2008).

40. R. Taft *et al.*, Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. *Nat. Struct. Mol. Biol.* **17**, 1030–1034 (2010).

41. C. Bracken *et al.*, Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic cleavage. *Nucleic Acids Research* **39**, 5658–5668 (2011).

42. E. L. Murray, D. R. Schoenberg, A+U-rich instability elements differentially activate 5'-3' and 3"-5" mRNA decay. *Mol. Cell. Biol.* **27**, 2791–2799 (2007).

43. I. Benilova, E. Karran, B. De Strooper, The toxic A[beta] oligomer and Alzheimer's disease: an emperor in need of clothes. *Nat Neurosci* **advance online publication** (2012), doi:10.1038/nn.3028.

44. G. G. Gross *et al.*, A. Lewin, Ed. Role of X11 and ubiquilin as in vivo regulators of the amyloid precursor protein in Drosophila. *PLoS ONE* **3**, e2495 (2008).

45. F. Riese *et al.*, R. Yan, Ed. Visualization and Quantification of APP Intracellular Domain-Mediated Nuclear Signaling by Bimolecular Fluorescence Complementation. *PLoS ONE* **8**, e76094 (2013).

Summary and Conclusion

sRNA profiles are altered by classical conditioning

The advent and widespread adoption of massively parallel nucleotide sequencing has pushed RNA to the forefront of modern biology. An ever growing catalog of previously unknown or underappreciated RNAs highlights how much remains to be understood concerning the fundamental mechanisms of life. sRNAs are the subject of intense study due to their important functions in gene regulation and the evolutionary conservation of the pathways they participate in. MicroRNAs have become a major area of research in biology generally, and within neuroscience in particular. Despite the increasing availability of massively parallel sequencing technology, most studies of microRNA function in neurons to date have focused on individual microRNAs. To our knowledge, no studies that examine genome wide microRNA expression in the context of *Drosophila* memory formation or behavior have been published. Significant efforts are also directed at understanding the functions of esiRNAs and piRNAs in somatic cells. Though a few published reports indicate that these classes of sRNAs also regulate important aspects of neural function and even behavior, such functions remain poorly studied.(*1*, *2*) Analyses of the now substantial set of publicly available sRNA sequencing data has also revealed a variety of other sRNA types aside from microRNAs, esiRNAs, and piRNAs that either have regulatory functions themselves, or that reveal the activity of other regulatory mechanisms.(3-6) Though much is now known of the basic mechanisms of sRNA function, how neurons utilize these mechanisms during memory formation

remains poorly understood. My efforts have been directed at identifying changes in gene regulation during the course of memory formation that involve sRNAs.

The literature now contains numerous instances in which individual microRNAs regulate aspects of synaptic plasticity and even memory in animals from nematodes to humans. However, microRNAs do not operate in isolation, and far fewer studies have examined the totality of microRNA involvement in these processes at once. Nonetheless, microRNAs are perhaps the best understood and most thoroughly cataloged class of sRNAs. Thus, a primary goal of my work was to profile genome wide microRNA expression in the *Drosophila* head, and to identify changes in this profile during memory formation. Here, I report in detail the expression of microRNA mapping sRNA reads in fly heads, and explore differences in expression between naïve flies, flies that have been subjected to olfactory aversive classical conditioning, and flies subjected to the conditioning protocol but exposed to only the US (Shock) or CS (Odor). I report expression levels for 316 of 427 known mature microRNAs, and show that 5 canonical mature microRNAs have significant changes in expression levels following LTM training. No microRNA was significantly regulated in the odor only condition, and only miR-958-5p changed significantly in the shock only condition. However, the odor only and shock only conditions each had only 3 samples, and additional replicates may yet show that some changes observed in these conditions are indeed significant. If not, the changes in microRNA expression that are significant only with the combination of US and CS may reflect important regulatory events involved in the integration

olfactory and noxious stimuli into a stabile memory. In general, changes observed in the shock only condition resemble those of LTM trained flies. This likely reflects the greater influence on microRNA expression induced by the activity of dopaminergic US circuitry than that by activity in the olfactory input to the MB. This makes sense conceptually, as the fly continually receives olfactory stimuli, but only stimuli associated with a US warrant remembering. Thus, it might be expected that US signaling would trigger changes in microRNA expression, though the absence of a salient CS with which to associate the US might trigger such changes in a variety of circuits that respond to many different stimuli. The overwhelming salience provided by the CS odor during LTM training might then refine US induced changes in microRNA expression through strong activation of circuits responding to the CS odor. Following this logic, the coordinated activation of specific circuits by the CS odor ensures that US induced changes in microRNA expression occur within a sufficiently large number of cells expressing a defined set of microRNAs to allow us to detect changes in that set. My observation that significant changes occur only in LTM trained flies might therefore persist even if more replicates were to be added to the shock only and odor only groups. While my results concerning distinctions between LTM conditioning and shock only or odor only treatments are less certain, the differences observed between LTM and control benefit from more replicates and are robust. Of note, I show that miR-312-3p is significantly downregulated following LTM training, and is the only microRNA significantly downregulated in any condition. miR-312-3p has been shown to regulate synaptic plasticity at the NMJ by controlling expression of the motor protein kinesin heavy chain.(7)

Therefore, my result fits with what is known regarding the function of miR-312-3p in neurons. However, the set of microRNAs I identify as being regulated following classical conditioning are together predicted to target transcripts of over 1000 genes. Beyond simply enumerating these targets, I hoped to gain insight into the regulatory output of microRNA modulation during memory formation. I present the results of a detailed target gene set analysis for regulated microRNAs, indicating that the G protein coupled signaling pathway is modulated through their activity. Of particular interest is the targeting of the dopamine receptor Dop1R1 by 4 out of 5 regulated microRNAs. That the US input into the MB is dopaminergic clearly raises the question of whether the significant changes in microRNA expression I report reflect feedback and feedforward mechanisms that tune dopamine response in the MB. As such, the relationship between LTM, changes in the expression of these microRNAs, and expression of Dop1R1 warrants investigation.

My analysis extends beyond canonical mature microRNAs, and documents the extensive presence of offset reads, occurrences of untemplated nucleotides, and tailing of microRNA reads. Such reads may reflect altered microRNA biogenesis, or the activity of mechanisms regulating the turnover of microRNAs.(8-10) Changes in the relative abundance of noncanonical microRNA reads might therefore reveal mechanisms coupling neural activity and microRNA production or turnover. Though noncanonical reads are common in my datasets, I do not observe statically significant changes in the relative abundance of any class for any microRNA. Thus, it

is not likely that microRNA processing or turnover are major avenues through which neural activity alters microRNA mediated silencing in *Drosophila*.

Recent work has demonstrated that the RNA editing enzyme ADAR influences differential expression of polycistronic microRNAs, and can alter microRNA processing and expression.(*11*, *12*) The miRISC associating protein FMRP interacts with and modulates the activity of ADAR in *Drosophila*.(13) ADAR is active in the nervous system of adult *Drosophila*, where it is required for normal behavior, and affects neural excitability.(*14*, *15*) These observations prompted me to examine my data for evidence of regulated microRNA editing during LTM formation. While I was able to detect what appear to be RNA editing events in microRNAs, the fraction of reads exhibiting editing at these sites did not change significantly in any of our experimental groups. Nonetheless, my analyses add to existing knowledge concerning noncanonical microRNAs, and the tissue specific nature of my data will be of use in future investigations of microRNA expression and editing in the *Drosophila* brain.

While my work provides a first genome wide glimpse of microRNA regulation during LTM formation in *Drosophila*, a number of questions will need to be addressed by further experimentation. As hoped, I was able to detect simultaneous changes in the expression levels of several microRNAs following LTM training. However, identifying which cells express each regulated microRNA will be a necessary step toward understanding how these changes together affect gene

expression. If regulated microRNAs are expressed in non-overlapping sets of cells, then expression of each microRNA's target genes can be considered independently. If some or all of the upregulated microRNAs are coexpressed, this may indicate that the set of target genes they share are subject to activity dependent silencing. If cells coexpress miR-312-3p and any of the upregulated microRNAs, interpretation may be more difficult. However, In this case, it may be fruitful to compare expression of genes targeted by miR-312-3p and upregulated microRNAs to those exclusively targeted by one or the other. As mentioned in chapter II, it may be worthwhile following training to compare expression of Dop1R1, which is targeted by miR-312-3p and 3 upregulated microRNAs, and Dop1R2 or Octβ3R, which are targeted only by miR-312-3p. Differential regulation of these genes could support a model in which changes in microRNA expression tune the MB response to US signaling following training.

My data and existing evidence from previously published studies clearly show that microRNAs participate in activity dependent gene regulation in neurons (Reviewed in (16)). Far less evidence is available in support of such a role for esiRNAs or piRNAs. This is in part due to their more recent discovery. It also stems from the greater number and diversity of esiRNA and piRNA sequences, as these sRNAs are produced from larger and less well defined regions of the genome. Indeed, while it is thought that most if not all expressed *Drosophila* microRNAs have now been identified and validated, esiRNA and piRNA expression remains poorly understood.(*17*, *18*) Moreover, how esiRNAs and piRNAs respond to neural activity

in *Drosophila* remains largely unknown. Mechanistically, it is also less clear how these sRNA classes might function in memory relevant gene regulation. It may be that they act by modulating heterochromatin. Alternatively, they might silence genes post-transcriptionally. To address such questions, I hoped to profile expression of esiRNAs and piRNAs in the *Drosophila* head, and to identify any changes in their expression following conditioning. I designed my sequencing library construction such that a size range of sRNAs encompassing esiRNAs and piRNAs would be captured in addition to microRNAs. Rather than solely relying on previously reported esiRNA and piRNA producing loci, I conducted computational analyses directed at identifying loci with features that are characteristic of these sRNA classes. In this way, I generated expression profiles for esiRNAs and piRNAs, and then conducted differential expression analysis for the loci I identified.

I found 368 esiRNA producing loci, some of which have been reported elsewhere, and some of which are novel.(*19*, *20*) 6 esiRNA loci display significant changes in read counts following LTM training. 4 of 6 regulated esiRNA loci correspond to a novel region of esiRNA production mapping to lysozyme family genes transcribed on both strands, and residing within an intron of the planar cell polarity (PCP) effector gene multiple wing hairs (mwh). Through analysis of reads mapping to multiple locations within this region, as well as through RNA secondary structure predictions, I show that these esiRNAs are likely derived from long hairpin structures formed by the mwh intron, and not from bidirectional transcription of lysozyme family genes. Lysozyme family genes participate in defense against

bacterial infection by catalyzing the hydrolysis of peptidoglycans present in bacterial cell walls, and it is therefore unclear how their regulation would be relevant to memory formation. However, mwh is a G-protein binding domain-formin homology 3 (GBD-FH3) protein in the frizzled pathway. It negatively regulates actin polymerization and filament formation, which are processes involved in cytoskeletal reorganization needed to produce structural plasticity at synaptic sites.(*21*, *22*) Thus, a model in which lysozyme family esiRNA production reflects regulation of mwh has far clearer logic than one in which lysozyme protein expression participates in memory formation. However, these ideas remain to be tested. Mwh has largely been studied in the context of PCP in the wing, and has not been directly implicated in synaptic plasticity or memory formation in *Drosophila*. It will therefore be important to evaluate the impact of mwh knockdown on neural function or memory. Such experiments could and should be designed such that the effects of perturbing the different splicing variants of mwh are evaluated. The TSS of mwh-RA is located between LysB and LysC, and may therefore be selectively silenced by production of lysozyme family esiRNAs. mwh-RB and mwh-RC have TSSs that are several kb upstream of LysB, and include the full lysozyme gene containing intron. Thus mwh-RA may be regulated differently from the other mwh variants. RNA hairpins directed against each mwh isoform could be conditionally expressed in specific brain regions at various time points around training. Mwh cDNA, which would presumably escape regulation by lysozyme family esiRNAs, could be similarly expressed in the brain following training. Such experiments could shed light on the function of mwh during memory formation.

piRNAs were first discovered and perhaps best characterized in *Drosophila*. The functions of piRNAs have been most extensively studied in the gonads, as the first phenotypes identified for piRNA pathway mutants related to gametogenesis and fertility. Early studies showed that piRNA pathways primarily function in defending the genome from selfish genetic elements in gametes (Reviewed in (23)). However, piRNAs are also produced in somatic cells, and can function as triggers for epigenetic silencing. Further, piRNAs that target the 3' UTRs of mRNAs have also been identified.(*24*, *25*) Recent publications now implicate piRNAs in control of genetic diversity within the *Drosophila* brain, and even in memory relevant transcriptional silencing.(*1*, *26*) I therefore sought to profile piRNA expression in *Drosophila* heads, and to identify piRNA loci that might respond to LTM training. I identified 82 likely piRNA loci, and conducted differential expression analysis on them. Many of the piRNA loci I identified have been previously reported, supporting the validity of the novel loci I identified. Regulated piRNA loci mapped to LTR retrotransposons, and in all instances exhibited decreased read counts in LTM trained flies vs. control, suggesting that learning might relieve these transposons from silencing via piRNAs. However, my RT-PCR experiments did not show any change in LTR retrotransposon levels. As piRNAs are largely derived from repetitive elements and transposons that are inserted at multiple locations throughout the genome, and tend to be expressed in abundance at these sites, my whole head lysate approach may have hampered my ability to detect changes only

occurring in circuits involved in olfactory memory. Methods permitting collection of piRNAs from specific cell types might overcome such obstacles in future work.

Identification of HECT sRNAs

The primary focus of my work was to identify changes in the expression of microRNAs, piRNAs, and esiRNAs during memory formation. However, novel classes of and functions for sRNAs continue to be discovered. These novel sRNAs can be functional, or merely reflect the activity of other regulatory mechanisms.(*4*, *6*, *27*, *28*) Such discoveries have often been the result of careful reanalysis of publicly available data sets, or from the examination of reads that are typically discarded during filtering of sRNA data. With this in mind, I have developed sequencing and analytical approaches that capture this type of information. In doing so, I sought to identify novel sRNAs and report their response to LTM training. My sequencing data includes many reads mapping to tRNAs, rRNAs, and other classes of known ncRNAs. The sheer volume of data returned from my sequencing experiments required me to focus my efforts elsewhere, but my preliminary examination suggests that tRNAs may produce sRNAs that are regulated by training. However, further exploration of this possibility remains for future work. My efforts toward identification of novel sRNAs lead me to describe HECT sRNAs as a class, and to report changes in read abundance for a subset of HECTs following LTM training. My work suggests that HECT sRNAs are degradative products from a subset of highly expressed mRNAs. What degradative mechanisms might drive HECT sRNA

production remain unclear, as do the reasons that HECTs are subject to such degradation while other transcripts are not. Degradome sequencing studies are shedding light on the intricacies of mRNA turnover, and could help answer these questions in the near future.(*29*, *30*) However, I found evidence that regulation of HECTs may be tied to the presence of a sequence element termed the cytoplasmic accumulation region element (CAR-E) that is known to increase accumulation and nuclear export of intronless transcripts. However, the significance of this finding remains unclear, and further experimentation will be required to determine if LTM induced changes in HECT expression are related to the presence of the CAR-E in transcripts. Mutation of the CAR-E within regulated HECTs would be relatively straightforward, and could reveal whether training induced changes in HECT expression are dependent on its presence.

dBACE is upregulated during LTM formation

My analysis of HECT sRNAs spurred a deeper analysis of dBACE expression following LTM training. dBACE has the largest increase in HECT sRNAs of any gene following LTM training. dBACE is also the homologue of human BACE genes, which are strongly implicated in Alzheimer's disease pathology. dBACE cleaves the amyloid precursor protein (APP) and its *Drosophila* homologue APPL such that toxic senile plaque forming peptides are liberated. However, senile plaque formation appears to occur only with aberrant APPL and/or dBACE expression. Recent work in mammals suggests that BACE activity is required for normal neurophysiology,

behavior, and memory.(*31*-*33*) APP family proteins and the products of their proteolytic processing affect neural morphology, synaptic plasticity, and behavior (Reviewed in (*34*, *35*)). Importantly, the highly conserved C-terminal fragment of APP family proteins is preferentially produced via β-γ-cleavage, and is a modulator of transcriptional programs.(*36*-*38*) My work produced evidence that increased dBACE expression following LTM training reflects a shift toward amyloidogenic processing of APPL, and leads to accumulation of AICD. These results are supported by behavioral experiments showing that dBACE and APPL expression are required for LTM, but not for STM or learning. Thus, the work presented here suggests strongly that LTM formation in flies involves a shift toward β-γ-processing of APPL, primarily driven by increased dBACE expression.

My efforts toward understanding the significance of HECT sRNA production, and related changes in dBACE expression have produced intriguing results, and prompt questions that will require additional experiments to address. Importantly, the identity of the APPL fragment that I suggest may be AICD must be confirmed. If the induced fragment is indeed AICD, this would suggest a line of questioning directed at understanding how AICD accumulation is coupled to synaptic plasticity and memory. It will also be critical to understand where in the brain dBACE is expressed, and in which cells its expression increases following training. Our data showing that dBACE and APPL expression in the MB is required for LTM suggests that such a search should begin in this brain region, but greater specificity could prove informative. Reporters of APPL cleavage and AICD nuclear localization

similar to those described in (*39*) could be used explore the transcriptional regulating activities of AICD during memory formation. In such experiments, nuclear localization of AICD would be revealed by the expression of a reporter, driven by the fusion of Gal4 to the C-terminus of APPL. Presumably, nuclear accumulation of AICD would largely be driven by β-secretase initiated APPL processing. This presumption could be checked through the use of dBACE mutants or knockdown.

Concluding Remarks

The original goal of my work was to understand how microRNA expression changes during memory formation. In designing experiments to tackle this issue, it became clear that a far greater variety of sRNAs could be sequenced, while still obtaining microRNA expression profiles. The unanticipated identification of training induced changes in lysozyme family esiRNAs and dBACE HECT sRNAs demonstrates the prudence of this approach. My work thus underscores the continuing value of survey type experiments in an era of abundant publicly available sequencing data. I had hoped to obtain a clear view of the contributions to sRNA regulation following conditioning from the CS and US separately, and in combination. However, the limited number of samples in the shock only and odor only conditions hampered this effort. Though the similarities between the shock only and LTM trained conditions suggests that most changes in sRNA expression

arise as a result of activity in US circuits, stronger claims in this regard will require additional samples in the shock only and odor only conditions.

The importance of microRNAs in controlling synaptic plasticity is now well established, though not fully understood. My work adds to what is known about microRNA participation in memory formation in that is, to my knowledge, the first simultaneous genome wide examination of microRNA expression during memory formation in *Drosophila*. However, the goal of such studies should be to examine expression of all microRNAs within particular neural circuits, or cells, rather than in whole head lysates. Further, experiments connecting genome wide changes in microRNA expression to changes in target gene expression, and ultimately to changes in synaptic plasticity will be necessary.

My examination of esiRNA expression during memory formation was productive in that it yielded the strongly regulated mwh/LysB-LysS esiRNA locus. However, that so few esiRNA loci display statistically significant changes in expression may support the view that esiRNAs are not a major mechanism for regulating synaptic plasticity. However, existing evidence is not sufficient to conclude one way or the other on this matter. As such, continued study of esiRNA involvement in memory formation may be warranted. Similarly, my work does not indicate that piRNAs are key players in memory formation. However, previous work in mollusks does indicate that piRNAs do regulate memory relevant gene expression.(1) Further, piRNAs do appear to have important functions in the Drosophila brain.(*26*, *40*) Thus, it is possible that piRNAs participate in *Drosophila* memory in ways that my work was unable to reveal.

My findings concerning regulation of dBACE during memory formation add to what is known about the mechanisms underlying AD. My work also bolsters *Drosophila* as a model organism with which to study AD. Indeed, *Drosophila* seems poised to make important contributions in this area. Recently, it was shown that cleavage of APPL by dBACE is required non-cell autonomously for glial survival in *Drosophila*.(*41*) This finding agrees with current research highlighting the contributions of glia to AD progression in mammals. (*42-44*) The power of *Drosophila* behavioral genetics could yield important mechanistic insight regarding the causes of AD. Moreover, the ease of work in flies, and the abundance of high-throughput *Drosophila* behavioral paradigms could make them a valuable system for use in drug development. β-secretase inhibitors have been a major area of effort for AD treatment. My work underscores the relevance of *Drosophila* for such studies, and supports the view that these interventions may not be the best approach for treating AD. In the future, *Drosophila* may prove valuable in efforts to understand and treat cognitive diseases such as AD. It is my hope that my work will be of use both in understanding AD, and in developing *Drosophila* as a model organism with which to study cognitive disease.

Literature cited

1. P. Rajasethupathy *et al.*, A Role for Neuronal piRNAs in the Epigenetic Control of Memory-Related Synaptic Plasticity. *Cell* **149**, 693–707 (2012).
2. B.-T. Juang *et al.*, Endogenous nuclear RNAi mediates behavioral adaptation to odor. *Cell* **154**, 1010–1022 (2013).
3. R. Taft *et al.*, Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. *Nat. Struct. Mol. Biol.* **17**, 1030–1034

(2010).

4. R. Taft, P. Hawkins, J. Mattick, K. Morris, The relationship between transcription initiation RNAs and CCCTC-binding factor (CTCF) localization. *Epigenetics & Chromatin* **4**, 13 (2011).

5. A. Sobala, G. Hutvagner, Small RNAs derived from the 5' end of tRNA can inhibit protein translation in human cells. *RNA Biol* **10**, 553–563 (2013).

6. modENCODE Consortium *et al.*, Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* **330**, 1787–1797 (2010).

7. K. Tsurudome *et al.*, The Drosophila miR-310 Cluster Negatively Regulates Synaptic Strength at the Neuromuscular Junction. *Neuron* **68**, 879–893 (2010).

8. S. L. Ameres, J.-H. Hung, J. Xu, Z. Weng, P. D. Zamore, Target RNA-directed tailing and trimming purifies the sorting of endo-siRNAs between the two Drosophila Argonaute proteins. *RNA (New York, N.Y.)* **17**, 54–63 (2011).

9. J. O. Westholm, E. Ladewig, K. Okamura, N. Robine, E. C. Lai, Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs. *RNA (New York, N.Y.)* **18**, 177–192 (2012).

10. R. Fukunaga *et al.*, Dicer partner proteins tune the length of mature miRNAs in flies and mammals. *Cell* **151**, 533–546 (2012).

11. G. Chawla, N. S. Sokol, ADAR mediates differential expression of polycistronic microRNAs. *Nucleic Acids Research* (2014), doi:10.1093/nar/gku145.

12. W. Yang *et al.*, Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nature Structural & Molecular Biology* **13**, 13–21 (2005).

13. B. Bhogal *et al.*, Modulation of dADAR-dependent RNA editing by the Drosophila fragile X mental retardation protein. *Nat Neurosci* **14**, 1517–1524 (2011).

14. X. Li, I. M. Overton, R. A. Baines, L. P. Keegan, M. A. O'Connell, The ADAR RNA editing enzyme controls neuronal excitability in Drosophila melanogaster. *Nucleic Acids Research* **42**, 1139–1151 (2014).

15. J. Jepson, R. Reenan, Adenosine-to-Inosine Genetic Recoding Is Required in the Adult Stage Nervous System for Coordinated Behavior in Drosophila. *J. Biol. Chem.* **284**, 31391–31400 (2009).

16. E. McNeill, D. Van Vactor, MicroRNAs Shape the Neuronal Landscape. *Neuron* **75**, 363–379 (2012).

17. X. Wang, S. Liu, Systematic Curation of miRBase Annotation Using Integrated Small RNA High-Throughput Sequencing Data for C. elegans and Drosophila. *Frontiers in Genetics* **2** (2011), doi:10.3389/fgene.2011.00025.

18. E. Berezikov *et al.*, Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. *Genome research* **21**, 203–215 (2011).

19. B. Czech *et al.*, An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**, 798–802 (2008).

20. K. Okamura *et al.*, The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**, 803–806 (2008).

21. W. J. Gault, P. Olguin, U. Weber, M. Mlodzik, Drosophila CK1-γ, gilgamesh, controls PCP-mediated morphogenesis through regulation of vesicle trafficking. *The Journal of Cell Biology* **196**, 605–621 (2012).

22. J. Yan *et al.*, The multiple-wing-hairs gene encodes a novel GBD-FH3 domain-containing protein that functions both prior to and after wing hair initiation. *Genetics* **180**, 219–228 (2008).

23. M. J. Luteijn, R. F. Ketting, PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nature Reviews Genetics* **14**, 523–534 (2013).

24. N. Robine *et al.*, A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr. Biol.* **19**, 2066–2076 (2009).

25. X. Huang *et al.*, A major epigenetic programming mechanism guided by piRNAs. *Developmental Cell* **24**, 502–516 (2013).

26. P. N. Perrat *et al.*, Transposition-driven genomic heterogeneity in the Drosophila brain. *Science* **340**, 91–95 (2013).

27. E. Berezikov *et al.*, Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. *Genome research* **21**, 203–215 (2011).

28. Z. Li *et al.*, Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Research* **40**, 6787–6799 (2012).

29. C. Bracken *et al.*, Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic

cleavage. *Nucleic Acids Research* **39**, 5658–5668 (2011).

30. P. Jackowiak, M. Nowacka, P. Strozycki, M. Figlerowicz, RNA degradome—its biogenesis and functions. *Nucleic Acids Research* **39**, 7361–7370 (2011).

31. I. Benilova, E. Karran, B. De Strooper, The toxic A[beta] oligomer and Alzheimer's disease: an emperor in need of clothes. *Nat Neurosci* **advance online publication** (2012), doi:10.1038/nn.3028.

32. D. Puzzo *et al.*, Endogenous amyloid-β is necessary for hippocampal synaptic plasticity and memory. *Annals of Neurology* **69**, 819–830 (2011).

33. F. M. Laird *et al.*, BACE1, a major determinant of selective vulnerability of the brain to amyloid-beta amyloidogenesis, is essential for cognitive, emotional, and synaptic functions. *The Journal of Neuroscience* **25**, 11693–11709 (2005).

34. K. T. Jacobsen, K. Iverfeldt, Amyloid precursor protein and its homologues: a family of proteolysis-dependent receptors. *Cell. Mol. Life Sci.* **66**, 2299–2318 (2009).

35. B. Poeck, R. Strauss, D. Kretzschmar, Analysis of amyloid precursor protein function in Drosophila melanogaster. *Exp Brain Res* **217**, 413–421 (2011).

36. C. Beckett, N. N. Nalivaeva, N. D. Belyaev, A. J. Turner, Nuclear signalling by membrane protein intracellular domains: The AICD enigma. *Cellular Signalling* **24**, 402–409 (2012).

37. B. B. Flammang *et al.*, Evidence that the amyloid-β protein precursor intracellular domain, AICD, derives from β-secretase-generated C-terminal fragment. *J Alzheimers Dis* **30**, 145–153 (2012).

38. F. Riese *et al.*, R. Yan, Ed. Visualization and Quantification of APP Intracellular Domain-Mediated Nuclear Signaling by Bimolecular Fluorescence Complementation. *PLoS ONE* **8**, e76094 (2013).

39. G. G. Gross *et al.*, A. Lewin, Ed. Role of X11 and ubiquilin as in vivo regulators of the amyloid precursor protein in Drosophila. *PLoS ONE* **3**, e2495 (2008).

40. A. Janic, L. Mendizabal, S. Llamazares, D. Rossell, C. Gonzalez, Ectopic expression of germline genes drives malignant brain tumor growth in Drosophila. *Science* **330**, 1824–1827 (2010).

41. B. J. Bolkan, T. Triphan, D. Kretzschmar, β-secretase cleavage of the fly amyloid precursor protein is required for glial survival. *The Journal of Neuroscience* **32**, 16181–16192 (2012).

42. S. Jo *et al.*, GABA from reactive astrocytes impairs memory in mouse models of

Alzheimer's disease. *Nature Medicine* **20**, 886–896.

43. K. V. Kuchibhotla, C. R. Lattarulo, B. T. Hyman, B. J. Bacskai, Synchronous Hyperactivity and Intercellular Calcium Waves in Astrocytes in Alzheimer Mice. *Science* **323**, 1211–1215 (2009).

44. B. A. Barres, The Mystery and Magic of Glia: A Perspective on Their Roles in Health and Disease. *Neuron* **60**, 430–440.

Appendix

Supplementary Figures

Figure S2.1. Mean normalized canonical miRNA counts

miRNA	CONTROL	LTM	ODOR	SHOCK
bantam-3p	269888.5882	247239.1294	325369.4314	293019.7451
bantam-5p	1806.843137	1957.905882	1368.215686	1153.254902
let-7-3p	28.3627451	17.95294118	14	26.15686275
let-7-5p	309420.3627	262324.6118	376999.0392	363086.4118
miR-1-3p	1995685.892	2003480.788	1851741.353	1851741.353
miR-1-5p	2.946078431	1.558823529	2.294117647	4.450980392
miR-10-3p	7796.303922	9919.682353	8627.72549	6699.352941
miR-10-5p	5675.323529	7560.682353	6427.411765	5229.215686
miR-100-3p	17.06372549	12.48823529	17.09803922	10.3627451
miR-100-5p	1952.372549	1501.235294	2022.411765	1739.941176
miR-1000-3p	99.48529412	119.8	90.6372549	83.42156863
miR-1000-5p	11138.79412	11271.92941	8991.901961	8849.509804
miR-1001-3p	12.14705882	13.79411765	21.30392157	15.68627451
miR-1001-5p	3277.166667	2795.647059	3148.666667	3769.098039
miR-1002-5p	1.710784314	1.911764706	1.450980392	1.862745098
miR-1003-3p	597.6960784	655.9058824	449.627451	596.0588235
miR-1003-5p	129.2205882	181.8588235	165.9411765	167.6568627
miR-1004-3p	213.0098039	312.9058824	159.6764706	201.5098039
miR-1005-3p	832.4607843	799.1411765	569.254902	609.0784314
miR-1006-3p	732.9411765	636.8470588	730.1372549	778.9607843
miR-1006-5p	49.53431373	78.04117647	36.95098039	32.01960784
miR-1007-3p	517.127451	293.5176471	647.1176471	582.9803922
miR-1007-5p	1.598039216	9.852941176	3.019607843	5.431372549
miR-1008-3p	74.65686275	70.72941176	50.32352941	66.85294118
miR-1009-3p	950.6568627	650.8705882	1136.588235	840.9803922
miR-1010-3p	9151.147059	10476.97647	10832.84314	9574.509804
miR-1010-5p	207.9215686	242.0705882	203.6862745	162.4215686
miR-1011-3p	35.53921569	44.34117647	37.03921569	38.17647059
miR-1012-3p	2926.696078	3159.329412	3637.921569	2605.588235
miR-1012-5p	1949.862745	2090.458824	1604.607843	1319.529412
miR-1013-3p	284.3333333	266.8	283.3333333	323.1568627
miR-1014-5p	1.12745098	1	1	1
miR-1015-3p	12.4754902	10.87058824	11.66666667	10.12745098
miR-1016-3p	5.352941176	4.841176471	4.676470588	6.539215686
miR-1016-5p	2.284313725	3.611764706	3.323529412	3.470588235
miR-1017-3p	328.1176471	350.4941176	282.9803922	240.6470588
miR-11-3p	30490.45098	31048.29412	36667.84314	35492.09804
miR-11-5p	424.3529412	266.0294118	312.9509804	369.3137255
miR-12-3p	50.89705882	59.56470588	59.31372549	74.84313725
miR-12-5p	17877.08824	18915.84706	13321.03922	14502.88235

Figure S2.1 (Continued)

miR-124-3p	3082.411765	2138.317647	3240.392157	3192.509804
miR-124-5p	299.5784314	247.7764706	260.4607843	306.8627451
miR-125-3p	221.2941176	206.3529412	142.4803922	155.3529412
miR-125-5p	14943.22549	16668.29412	12903.84314	10783.19608
miR-133-3p	1488.196078	1321.188235	1143.980392	1122.54902
miR-133-5p	158.9754902	173.9470588	113.6372549	102.9019608
miR-137-3p	7772.245098	5412.670588	7624.098039	7564.019608
miR-137-5p	19.89215686	18.91764706	19.02941176	14.40196078
miR-13a-3p	290.5294118	206.3470588	298.8921569	328.372549
miR-13a-5p	26.86764706	9.5	23.38235294	46.02941176
miR-13b-3p	3849.019608	3897.364706	2922.098039	2645.137255
miR-13b-5p	43.78921569	28.58235294	37.64705882	32.40196078
miR-14-3p	30708.79412	28182.16471	40402.11765	32289.01961
miR-14-5p	2400.401961	1895.394118	4965.529412	3585.431373
miR-184-3p	2571464.049	2369941.694	2715408.588	2715408.588
miR-184-5p	36.38235294	40.98235294	31.11764706	26.44117647
miR-190-3p	6.362745098	4.023529412	5.588235294	4.882352941
miR-190-5p	8708.5	6099.529412	8375.54902	9725.72549
miR-193-3p	206.8137255	133.9647059	199.6960784	218.2156863
miR-193-5p	106518.5196	110043.5059	90377.17647	84208.58824
miR-210-3p	40945.17647	43440.24706	34552.76471	33130.56863
miR-210-5p	293.8137255	282.2588235	259.745098	241.5098039
miR-219-3p	47.10784314	44.55882353	109.2843137	93.81372549
miR-219-5p	210.8235294	83.04705882	406.5882353	454.7843137
miR-2279-3p	3.166666667	3.111764706	6.12745098	5.62745098
miR-2279-5p	22.13235294	13.54705882	19.15686275	33.30392157
miR-2282-3p	11.93137255	6.982352941	10.08823529	12.6372549
miR-2283-3p	3.691176471	1	4.068627451	1
miR-2283-5p	1.401960784	1	2.039215686	1.470588235
miR-2489-3p	375.3333333	297.1176471	277.5490196	344.2156863
miR-2490-5p	3.264705882	3.358823529	1.803921569	3.470588235
miR-2491-3p	1.225490196	1	1	1
miR-2491-5p	1.220588235	1.5	1	1
miR-2492-3p	1.411764706	1	3.303921569	1.431372549
miR-2492-5p	1.117647059	1	1.529411765	1
miR-2494-3p	1.220588235	1.258823529	1.450980392	1.431372549
miR-2496-3p	3.362745098	2.305882353	2.254901961	2.745098039
miR-2496-5p	1.318627451	2.005882353	2.196078431	2.803921569
miR-2497-3p	3.176470588	2.017647059	4.049019608	1
miR-2497-5p	1.651960784	1	1	1
miR-2499-3p	1	1.188235294	1	1
miR-2500-3p	58.59803922	60.22352941	39.66666667	46.16666667

Figure S2.1 (Continued)

miR-2500-5p	11.20098039	13.07058824	5.578431373	12.17647059
miR-2501-3p	1	1	1	1
miR-2501-5p	2.43627451	2.464705882	3.588235294	3.431372549
miR-252-3p	24.17156863	21.27647059	21.08823529	18.7745098
miR-252-5p	61393.30392	54307.17647	67713.07843	52283.17647
miR-2535b-3p	222.7156863	237.2705882	270.4509804	252.8627451
miR-263a-3p	54.42156863	51.92941176	53.3627451	69.17647059
miR-263a-5p	694123.7745	850055.1765	883104.0588	802718.7647
miR-263b-3p	27.34803922	42.13529412	13.23529412	18.64705882
miR-263b-5p	182405.8725	230427.3176	227987.5686	227987.5686
miR-274-3p	2.029411765	1.3	2.411764706	1
miR-274-5p	33074.0098	30223.70588	33611.2549	32909.52941
miR-275-3p	18434.06863	19709.95294	14208.66667	11670.62745
miR-275-5p	63.40196078	61.28823529	49.33333333	40.54901961
miR-276a-3p	343515.1471	311215.6941	195637.8824	225696.098
miR-276a-5p	795.0882353	721.8058824	808.1666667	929.1862745
miR-276b-3p	58134.47059	45172.36471	50924.4902	58941.31373
miR-276b-5p	795.0882353	721.8058824	808.1666667	929.1862745
miR-277-3p	27744.10784	31049.62353	24599.05882	25123.35294
miR-277-5p	432.2254902	391.1647059	397.3627451	373.8431373
miR-278-3p	2266.941176	1340.105882	3220.411765	2886.901961
miR-278-5p	3720.607843	2150.276471	9258.901961	5221.72549
miR-279-3p	5503.127451	5218.047059	5433.235294	5882.784314
miR-279-5p	2.431372549	2.664705882	2.607843137	1
miR-281-3p	1195.083333	1200.6	1532.470588	1273.235294
miR-281-5p	20.5	18.71764706	34.99019608	29.31372549
miR-282-3p	1753.441176	1572.282353	1336.470588	1447.529412
miR-282-5p	543.9019608	228.3705882	688.0784314	896.8627451
miR-283-3p	70.74509804	85	68.58823529	56.53921569
miR-283-5p	7270.833333	7987.341176	7039.588235	8488.215686
miR-284-3p	1.357843137	1	1	1
miR-284-5p	10105.55882	10222.91765	9835.607843	8898.764706
miR-285-3p	2746.313725	2599.247059	2306.352941	2091.980392
miR-285-5p	8039.362745	9449.517647	9842.156863	8746.882353
miR-286-3p	86.76470588	78.74117647	104.6862745	106.9901961
miR-286-5p	1	1.194117647	1	1.901960784
miR-2a-3p	11483.29412	11069.70588	6968.176471	7525.098039
miR-2a-5p	49.59803922	33.51176471	59.30392157	57.1372549
miR-2b-3p	7414.666667	7309.952941	6643.352941	5662.803922
miR-2b-5p	22.89705882	28.57058824	20.3627451	19.83333333
miR-2c-3p	1235.235294	1394.364706	859.4509804	866.1176471
miR-2c-5p	1421.764706	1196.282353	2189.176471	1768.862745

Figure S2.1 (Continued)

miR-3-3p	3.137254902	1.105882353	2.803921569	4.519607843
miR-303-3p	1.240196078	1	1	1
miR-304-3p	18.66176471	18.31764706	14.79411765	12.93137255
miR-304-5p	3219.22549	2723.882353	3986.529412	4209.254902
miR-305-3p	3303.421569	3745.4	2440.960784	2156.882353
miR-305-5p	12786.78431	10331.49412	10390.43137	12706.45098
miR-306-3p	316.5784314	310.2588235	244	200.2843137
miR-306-5p	2625.764706	904.0941176	3177.215686	3717.490196
miR-307a-3p	7328.441176	7120.258824	5340.352941	6171
miR-307a-5p	569.1078431	702.1411765	372.4901961	560.8039216
miR-307b-3p	25.30882353	22.72352941	17.46078431	20.21568627
miR-307b-5p	9.12745098	6.764705882	3.637254902	2.637254902
miR-308-3p	198.9411765	152.8294118	211.8431373	327.4901961
miR-308-5p	65.25490196	14.34117647	61.40196078	110.9019608
miR-310-3p	1.367647059	1.105882353	1	1
miR-310-5p	1.617647059	1.429411765	1	1
miR-311-3p	20.03921569	10.00588235	16.2254902	8.058823529
miR-311-5p	1.617647059	1.241176471	1	1
miR-312-3p	180.377451	41.90588235	77.54901961	63.64705882
miR-312-5p	1	1.241176471	1	1
miR-313-3p	1.740196078	1	1	1
miR-313-5p	1.31372549	1	1.450980392	1.431372549
miR-314-3p	997.1470588	2452.635294	1329.705882	2701.784314
miR-314-5p	14.8872549	36.41764706	29.45098039	65.67647059
miR-315-3p	21.70098039	12.8	15.51960784	30.12745098
miR-315-5p	39317.79412	31872.94118	40866.37255	42066.23529
miR-316-3p	320.627451	143.5647059	786.3529412	794.3137255
miR-316-5p	1669.132353	1603.082353	1989.352941	2183.411765
miR-317-3p	665699.8333	752790.0706	443856.6275	520425
miR-317-5p	294.1372549	286.7647059	239.9803922	240.1960784
miR-318-3p	135.4019608	49.14705882	47.97058824	54.41176471
miR-318-5p	1.098039216	1	1	1
miR-31a-3p	20.93627451	23.36470588	14.74509804	14.82352941
miR-31a-5p	8185.754902	6890.623529	17307.62745	10114.13725
miR-31b-5p	294.9019608	331.3294118	305.0392157	330.0784314
miR-33-3p	151.3088235	149.0941176	120.9607843	150.127451
miR-33-5p	15429.79412	15969.95294	13147.27451	15074.88235
miR-34-3p	2231.352941	1952.905882	1581.215686	1250.921569
miR-34-5p	33071.82353	45512.77647	25863.72549	27768.58824
miR-3641-5p	1.549019608	1.105882353	2.294117647	2.37254902
miR-3642-3p	1.093137255	1.241176471	1	1
miR-3642-5p	1.833333333	1.970588235	1.450980392	2.333333333

Figure S2.1 (Continued)

miR-3643-3p	1	1.194117647	1	1
miR-3643-5p	1.220588235	1.264705882	1.450980392	1.431372549
miR-3644-3p	1	1.241176471	1	1
miR-3645-3p	2.740196078	2.547058824	4.333333333	2.333333333
miR-3645-5p	3.392156863	2.629411765	3.098039216	4.343137255
miR-375-3p	5032.22549	4334.2	4893.647059	5223.980392
miR-375-5p	256.9117647	436.5529412	291.1176471	283.8235294
miR-4-3p	2.534313725	4.052941176	1	2.284313725
miR-4-5p	1	1	1	1
miR-4908-3p	1.367647059	1	1	1
miR-4910-5p	4.411764706	3.535294118	3.382352941	3.87254902
miR-4911-3p	1	1	1	1
miR-4912-5p	1.117647059	1.447058824	1	1.901960784
miR-4913-3p	16.39705882	10.55882353	11.2254902	16.97058824
miR-4914-5p	1.406862745	1	1	1
miR-4915-5p	2.259803922	4.029411765	2.31372549	1.901960784
miR-4916-3p	31.47058824	16.45882353	24.8627451	35.01960784
miR-4919-5p	21.28431373	24.60588235	9.607843137	7.5
miR-4939-3p	1	1	1	1
miR-4940-3p	4.700980392	4.423529412	10.52941176	4.343137255
miR-4940-5p	19.24019608	20.42352941	29.42156863	30.70588235
miR-4941-5p	1	1.258823529	1	1
miR-4942-3p	2.401960784	4.170588235	8.039215686	2.637254902
miR-4943-3p	4.848039216	10.23529412	12.19607843	6.470588235
miR-4946-5p	2.848039216	1.264705882	1	1
miR-4947-5p	1	1.241176471	1	1.901960784
miR-4949-3p	1.62745098	1	1.450980392	1.431372549
miR-4949-5p	1.892156863	1.982352941	1	1
miR-4950-3p	1.098039216	1.741176471	1.450980392	1.431372549
miR-4951-3p	4.338235294	3.941176471	2.137254902	6.107843137
miR-4951-5p	589.4607843	559.7882353	579.6470588	505.2352941
miR-4952-3p	5.31372549	6.294117647	3.058823529	2.696078431
miR-4952-5p	39.62254902	35.08823529	38.06862745	32.90196078
miR-4955-3p	1.357843137	1.794117647	2.31372549	3.549019608
miR-4956-3p	6.568627451	5.623529412	9.921568627	5.549019608
miR-4956-5p	1.465686275	1	1	1
miR-4957-3p	2.230392157	2.470588235	2.529411765	1.431372549
miR-4957-5p	1.705882353	2.341176471	1.745098039	5.401960784
miR-4958-5p	2.254901961	2.388235294	1.294117647	1.81372549
miR-4959-5p	1.215686275	1.188235294	1	1
miR-4960-3p	210.2156863	232.9529412	241.2745098	203.5098039
miR-4961-3p	5.019607843	2.129411765	7.137254902	1

Figure S2.1 (Continued)

miR-4961-5p	3.470588235	3.705882353	4.156862745	9.578431373
miR-4962-3p	12.84313725	14.67647059	34.44117647	39.80392157
miR-4962-5p	1	1	1.294117647	1
miR-4963-3p	6.754901961	6.2	7.519607843	9.843137255
miR-4964-3p	2.681372549	3.352941176	3.764705882	1
miR-4965-5p	1	1	1	1.901960784
miR-4966-3p	1	1	3.029411765	1
miR-4966-5p	1	1	1	1.882352941
miR-4968-5p	2.235294118	2.347058824	3.607843137	1
miR-4969-5p	242.3529412	238.4588235	326.1764706	188.9803922
miR-4971-5p	5.651960784	8.352941176	7.107843137	3.205882353
miR-4972-3p	1	1	1.529411765	1
miR-4972-5p	1	1	1	1
miR-4973-3p	3.593137255	2.788235294	1.803921569	1.431372549
miR-4973-5p	14.51470588	24.95294118	8.803921569	8.176470588
miR-4974-3p	2.401960784	1.264705882	1.450980392	1
miR-4974-5p	1	1	1.450980392	1
miR-4975-5p	3.799019608	4.011764706	5.117647059	5.196078431
miR-4976-5p	7.848039216	5.923529412	2.470588235	10
miR-4977-3p	5.137254902	7.082352941	5.245098039	7.245098039
miR-4978-5p	1.093137255	1	1	1
miR-4979-5p	1.240196078	1	1	1
miR-4980-3p	1.460784314	1.3	1.450980392	1.431372549
miR-4981-3p	1.632352941	2.411764706	1.529411765	1.901960784
miR-4982-3p	1.338235294	1	1.745098039	1
miR-4982-5p	2.058823529	1	1	4.450980392
miR-4983-3p	3.25	2.270588235	3.833333333	1.901960784
miR-4983-5p	1.215686275	1.723529412	1.745098039	2.696078431
miR-4984-3p	4.774509804	6.147058824	2.31372549	5.401960784
miR-4985-3p	1	1.194117647	1	1
miR-4985-5p	2.019607843	3.811764706	1.745098039	2.637254902
miR-4987-3p	1.098039216	1.723529412	1.745098039	1.431372549
miR-5-3p	1.240196078	2.570588235	1.509803922	2.725490196
miR-5-5p	66.77941176	57.55882353	66.7745098	78.80392157
miR-6-3p	4.475490196	5.347058824	6.205882353	7.254901961
miR-6-5p	1.62745098	1.5	1.529411765	1
miR-7-3p	21.84803922	13.98823529	26.67647059	26.40196078
miR-7-5p	32133.58824	36601.18824	56724.98039	32061.17647
miR-79-3p	567.5686275	425.5058824	625.3137255	782.9019608
miR-79-5p	113.7598039	97.27058824	97.02941176	104.7352941
miR-8-3p	475534.1078	516768.0471	472690	472690
miR-8-5p	75451.55882	94960.6	79513.09804	84893.92157

Figure S2.1 (Continued)

miR-87-3p	6998.088235	8694.647059	5280.117647	4679.784314
miR-87-5p	100.0833333	130.3176471	95.90196078	78.58823529
miR-927-3p	11275.22549	10322.94118	14784.03922	13672.60784
miR-927-5p	18099.2451	15007.85882	16780.5098	18550.86275
miR-929-3p	572.3039216	397.4117647	574.5490196	669.2352941
miR-929-5p	6247.980392	5004.258824	5652.039216	7010.647059
miR-92a-3p	57.75980392	57.57647059	54.54901961	43.56862745
miR-92a-5p	396.4019608	505.1176471	481.6666667	396.4117647
miR-92b-3p	354.4460784	406.1411765	399	294.0882353
miR-92b-5p	1	1.258823529	1	1
miR-932-3p	201.5147059	201.5882353	259.8823529	165.6078431
miR-932-5p	20411.69608	14287.11765	18935.17647	19626.11765
miR-954-3p	1.794117647	2.505882353	1.803921569	2.696078431
miR-954-5p	251.0196078	310.6117647	248.254902	207.3137255
miR-955-3p	7.421568627	9.064705882	10.3627451	5.607843137
miR-955-5p	13.76470588	13.91176471	5.578431373	5.382352941
miR-956-3p	9160.323529	19173.57647	15355.17647	33107.5098
miR-956-5p	49.44117647	143.5117647	91.39215686	249.2352941
miR-957-3p	254726.8039	229169.0941	262903.0196	279107.1176
miR-957-5p	55.68627451	49.29411765	60.61764706	39.6372549
miR-958-3p	735.6666667	2945.164706	973.0784314	1975.960784
miR-958-5p	360.9901961	1213.647059	765.3921569	1519.490196
miR-959-3p	13.23529412	13.95882353	13.65686275	7.794117647
miR-959-5p	8.225490196	7.3	6.735294118	6.343137255
miR-960-3p	8.617647059	12.31764706	3.460784314	5.960784314
miR-960-5p	12.18627451	23.99411765	18.23529412	19.23529412
miR-961-3p	1.225490196	2.123529412	1	1.862745098
miR-961-5p	2.965686275	3	1.803921569	2.803921569
miR-962-5p	2.362745098	3.347058824	1.529411765	1.431372549
miR-963-3p	1.12745098	1	1	1
miR-963-5p	1.519607843	1.723529412	1.803921569	1
miR-964-3p	1.607843137	1	1	1.431372549
miR-964-5p	4.06372549	4.982352941	5.882352941	6.81372549
miR-965-3p	29.42647059	25.44117647	28.2745098	50.73529412
miR-965-5p	2482.45098	2453.694118	2022.843137	1566.215686
miR-966-3p	1.117647059	1.435294118	1	1
miR-966-5p	139.7058824	158.9764706	173.254902	154.1372549
miR-967-3p	1	1.970588235	1	2.284313725
miR-967-5p	128.5	53.88235294	233.7941176	191.8431373
miR-968-5p	1.191176471	1.264705882	1.450980392	1.431372549
miR-969-3p	77.87254902	69.34117647	60.62745098	64.87254902
miR-969-5p	11.76470588	17.1	5.411764706	11.06862745

Figure S2.1 (Continued)

miR-970-3p	5603.833333	7218.529412	3799.901961	3751.392157
miR-970-5p	99.5245098	129.9176471	133.2941176	109.7745098
miR-971-3p	37.55392157	43.62941176	57.26470588	32.97058824
miR-971-5p	119.2156863	127.5294118	88.54901961	77.64705882
miR-972-3p	15.78921569	19.85882353	32.6372549	14.44117647
miR-973-3p	1.946078431	1.905882353	2.529411765	1.901960784
miR-973-5p	2.421568627	2.729411765	1.901960784	1
miR-974-5p	1.220588235	1.623529412	4.784313725	1.862745098
miR-975-5p	1.529411765	1.3	1.450980392	1.431372549
miR-976-3p	17.78921569	14.77647059	8.264705882	8.519607843
miR-977-3p	1.828431373	1.729411765	2.921568627	1.901960784
miR-977-5p	1.705882353	1.452941176	1	1
miR-978-3p	3.485294118	2.594117647	4.980392157	5.490196078
miR-978-5p	1	1.194117647	1	1
miR-979-3p	1	1.105882353	1.529411765	2.294117647
miR-979-5p	1.12745098	1	1	2.31372549
miR-980-3p	16.84803922	7.511764706	15.7254902	20.67647059
miR-980-5p	423	491.6823529	463	414.5882353
miR-981-3p	13124.55882	9919.082353	11153.35294	11725.35294
miR-981-5p	50.83333333	59.32941176	43.12745098	53.44117647
miR-982-3p	1	1	1	2.764705882
miR-982-5p	50.43137255	60.35882353	58.18627451	48.76470588
miR-983-3p	1.093137255	1.294117647	2.254901961	1.862745098
miR-983-5p	183.745098	152.7176471	195.2156863	213.3333333
miR-984-3p	1.519607843	1	1.529411765	1
miR-984-5p	2.083333333	3.023529412	3.588235294	2.294117647
miR-985-3p	1.950980392	1.5	1	1
miR-986-3p	4.289215686	2.782352941	5.392156863	6.029411765
miR-986-5p	4.289215686	2.129411765	3.460784314	6.343137255
miR-987-3p	53.04901961	52.44705882	30.74509804	23.56862745
miR-987-5p	52767.4902	53998.37647	51946.80392	51495.41176
miR-988-3p	4269.186275	3619.223529	2441.078431	2185.823529
miR-988-5p	1878.303922	1642.141176	1548.176471	1426.568627
miR-989-3p	252.1911765	43.72941176	89.40196078	56.35294118
miR-990-3p	3.931372549	2.647058824	2.558823529	3.470588235
miR-990-5p	501.4607843	575.4352941	497.2941176	461.3921569
miR-991-3p	2.931372549	2.411764706	2.852941176	3.950980392
miR-992-3p	1	1	1	1
miR-993-3p	587.0588235	517.2117647	630.5686275	345
miR-993-5p	176.1421569	190.1647059	184.0098039	162.372549
miR-994-3p	1.691176471	1	3.588235294	3.480392157
miR-994-5p	21.38235294	2.223529412	11.1372549	2.245098039

Figure S2.1 (Continued)

miR-995-3p	463.5686275	521.0588235	353.1568627	436.5882353
miR-995-5p	45.08823529	40.93529412	28.06862745	25.19607843
miR-996-3p	3538.078431	2966.552941	3319.098039	4547.156863
miR-996-5p	497.754902	586.0352941	427.4705882	516
miR-998-3p	530.1764706	567.4705882	473.7647059	490.1764706
miR-998-5p	1435.539216	1119.670588	1164.078431	1050.882353
miR-999-3p	57430.87255	30981.56471	67990.92157	74947.27451
miR-999-5p	15.73039216	18.65294118	13.25490196	13.81372549
miR-9a-3p	76.34803922	75.74117647	83.82352941	73.24509804
miR-9a-5p	60806.7549	59099.83529	62831.60784	86406.7451
miR-9b-3p	639.5784314	637.5058824	488.0588235	481.9411765
miR-9b-5p	5425.529412	4859.094118	4516.352941	5527.235294
miR-9c-3p	3.745098039	2.617647059	6.068627451	10.2254902
miR-9c-5p	61227.17647	69701.94118	49921.11765	55877.29412
miR-iab-4-5p	22.57352941	10.71176471	42.15686275	15.82352941
miR-iab-8-5p	4.642156863	2.817647059	4.882352941	1.901960784

Canonical microRNA reads were counted for each library. These values were normalized using the quantile method of the limma Bioconductor package.(*1*) Normalized read counts were then averaged within each condition.

Figure S2.2 Differential expression analysis for canonical miRNAs

		LTM vs. Control			Odor vs. Control			Shock vs. Control		
miRNA	logCPM	logFC	PValue	FDR	logFC	PValue	FDR	logFC	PValue	FDR
bantam-3p	15.23585737	-0.058641294	0.880146895	1	0.273408894	0.570653575	1	-0.032526346	0.946404969	1
bantam-5p	7.748545166	0.099950208	0.818797211	1	-0.300485769	0.585134712	1	-0.585055405	0.288202159	0.936915813
let-7-3p	1.685256517	-0.578969888	0.291292912	1	-1.289216401	0.075772142	1	-0.542052855	0.433739651	1
let-7-5p	14.65036861	0.153503373	0.688802168	1	-0.049012437	0.92491275	1	-0.330208098	0.532671216	1
miR-1-3p	18.28477604	0.330378844	0.51708143	1	0.030615613	1	1	-0.331340128	0.333907745	1
miR-10-3p	9.635691567	0.244281423	0.562583455	1	0.087128527	0.869317418	1	-0.039922339	0.939414573	1
miR-10-5p	8.712946695	0.518800502	0.204711108	1	0.337279151	0.513043657	1	0.154533213	0.76577772	1
miR-100-5p	8.107038715	-0.204086467	0.630663781	1	-0.459303405	0.393002767	1	-0.543708535	0.313811019	1
miR-1000-3p	3.73856449	0.141021582	0.770623235	1	-0.104405035	0.862542529	1	-0.416024234	0.493700551	1
miR-1000-5p	9.426926455	0.110611297	0.786898747	1	-0.164019215	0.748577324	1	-0.289947021	0.571550241	1
miR-1001-5p	2.554266029	-0.57406814	0.280446351	1	-0.815268596	0.23460798	1	-0.927554277	0.168795431	1
miR-1003-3p	6.607792263	0.224754893	0.611693062	1	-0.163545309	0.770820322	1	0.07739698	0.888438459	1
miR-1004-3p	5.378139854	0.330525642	0.47500414	1	-0.335100477	0.570030178	1	-0.086727602	0.879477299	1
miR-1005-3p	7.265455106	-0.152635307	0.73227475	1	-0.486778394	0.401630996	1	-0.545551959	0.339866723	1
miR-1006-3p	7.067273225	-0.172446262	0.688336684	1	0.063577008	0.905390947	1	-0.02590386	0.961230673	1
miR-1007-3p	4.216328839	-1.032799917	0.033862413	0.501779388	-0.714916274	0.280451393	1	-0.846293431	0.213563485	1
miR-1009-3p	7.239357425	-0.382926733	0.36600589	1	0.145004747	0.784570492	1	-0.045450692	0.933527356	1
miR-1010-3p	10.62132796	-0.020360865	0.960892584	1	-0.197465523	0.696858631	1	-0.384481711	0.448443559	1
miR-1011-3p	2.682050373	0.401221925	0.423291755	1	-0.04780464	0.939967688	1	0.161616628	0.79467726	1
miR-1012-3p	9.273458653	-0.069019624	0.868611295	1	0.099402241	0.846092447	1	-0.38147816	0.45851096	1
miR-1017-3p	5.208223806	-0.084661172	0.853331791	1	-0.19990445	0.729791681	1	-0.584863983	0.311466824	1
miR-11-3p	12.09822583	-0.174423621	0.666587895	1	-0.219256108	0.660904305	1	-0.42432024	0.397301287	1
miR-11-5p	5.519117565	-0.43976099	0.319838071	1	-0.621314317	0.279155453	1	-0.671593258	0.241421439	1
miR-12-3p	3.501241861	0.223367693	0.639491591	1	0.400112499	0.509226225	1	0.557129005	0.352114284	1
miR-12-5p	10.89944788	0.102212953	0.814337535	1	-0.661650428	0.244770336	1	-0.456950407	0.425242237	1
miR-124-3p	7.267934818	-0.101714527	0.809850565	1	0.098257782	0.853978653	1	-0.144910433	0.786148493	1
miR-124-5p	5.399853181	-0.021541978	0.961434188	1	-0.303392888	0.589604817	1	-0.254123281	0.65105351	1
miR-125-3p	4.698702483	-0.339145188	0.475267596	1	-0.688202529	0.264204034	1	-0.561414236	0.356376048	1
miR-125-5p	11.28574705	-0.037567174	0.925230477	1	0.010880454	0.982720545	1	-0.286649132	0.568966393	1
miR-133-3p	7.449865748	-0.000846242	0.998451187	1	-0.477885449	0.383441468	1	-0.574396948	0.301261071	1
miR-133-5p	3.841850651	0.014216382	0.977784559	1	-0.504979788	0.431559153	1	-0.675069985	0.295330611	1
miR-137-3p	10.43541844	-0.470830186	0.247289438	1	-0.301468712	0.54850661	1	-0.407454375	0.418650447	1
miR-137-5p	1.595784662	-0.294348005	0.598366234	1	0.028197966	0.968776516	1	-0.702717172	0.327851782	1
miR-13a-3p	3.533710139	-0.103707445	0.833821927	1	-0.254637777	0.689037917	1	-0.140466199	0.826325974	1
miR-13b-1-5p	2.413520346	-0.529332729	0.304892713	1	-0.137752014	0.827448767	1	-0.569765251	0.384402643	1
miR-13b-2-5p	2.612289157	-0.24032295	0.639420172	1	-0.323361543	0.620225516	1	0.001110794	0.999726672	1
miR-13b-3p	8.725166527	0.06207687	0.881736551	1	-0.088653213	0.864723321	1	-0.331061588	0.526400705	1
miR-14-3p	12.0537345	-0.001839887	0.996454858	1	-0.091346829	0.85494692	1	-0.190125855	0.719348914	1
miR-14-5p	8.041557128	-0.115341883	0.780109392	1	0.002502451	0.996945078	1	-0.492755038	0.390688238	1
miR-184-3p	18.32086626	-0.256756256	0.511246169	1	-0.197257539	0.041560447	1	-0.619105305	0.125966855	1
miR-190-5p	7.389700671	-0.586467432	0.173556832	1	-0.287675342	0.586193395	1	-0.394142478	0.466183792	1
miR-193-3p	3.998883668	-0.072517416	0.879788213	1	-0.374011105	0.553586279	1	-0.347879403	0.575747906	1
miR-193-5p	13.64050009	-0.000274807	0.9993906	1	-0.284372923	0.564620292	1	-0.371495782	0.452486724	1
miR-210-3p	8.609659059	-0.099246989	0.809411624	1	0.229880784	0.659101942	1	0.174973087	0.735440739	1
miR-210-5p	2.979872847	-0.324538082	0.522200907	1	-0.615728315	0.336402279	1	-0.837964068	0.207128718	1
miR-219-3p	2.727046757	0.143711284	0.775061965	1	0.432530256	0.536585761	1	0.137144452	0.829226147	1
miR-219-5p	3.740475205	-1.075333531	0.049378954	0.670730793	-1.211367214	0.106051779	1	-1.30303325	0.092974656	1
miR-2489-3p	5.086020548	-0.065999343	0.883817158	1	-0.182716768	0.746438829	1	-0.403121472	0.476456102	1
miR-2500-3p	2.494822817	0.093567188	0.853239418	1	-0.145622992	0.825825878	1	-0.212857923	0.74743275	1
miR-252-5p	13.14643863	-0.018054799	0.962746767	1	-0.087402212	0.864693773	1	-0.434726706	0.392183215	1
miR-263a-3p	2.53231484	-0.276078887	0.583264591	1	-0.207333025	0.754198372	1	-0.007177753	0.990358756	1
miR-263b-3p	1.969115731	0.523282952	0.338984793	1	-1.083712282	0.145291602	1	-0.364217521	0.61124957	1
miR-263b-5p	13.9697347	0.127116337	0.745528884	1	-0.194612296	0.70916219	1	-0.495722364	0.34577413	1
miR-275-3p	10.79453277	0.298617439	0.477758016	1	-0.270670149	0.608148086	1	-0.604628504	0.259781089	1
miR-276a-3p	15.51528667	-0.137284434	0.726270676	1	-0.339176352	0.482972245	1	-0.374603187	0.439379051	1
miR-276b-3p	12.93522605	-0.246854625	0.537350539	1	-0.514966907	0.309569348	1	-0.421847142	0.407461915	1
miR-277-3p	11.69884567	-0.099867638	0.798984309	1	0.437490022	0.388771568	1	0.230811686	0.650269337	1
miR-277-5p	5.378827353	-0.184264525	0.676340536	1	0.135003915	0.812340512	1	-0.226730581	0.68610722	1
miR-278-3p	7.412567753	-0.493802602	0.268264952	1	-0.798634538	0.191235299	1	-0.83923715	0.173240519	1
miR-278-5p	9.073283461	-0.487447906	0.238756606	1	0.070951448	0.902035124	1	-0.375823906	0.49601967	1
miR-279-3p	6.795250803	0.011010601	0.979439898	1	0.084076329	0.878259073	1	0.196950035	0.725347188	1
miR-281-2-5p	6.097648577	0.075198734	0.862628118	1	0.499594196	0.359104336	1	0.186903333	0.729179235	1

Figure S2.2 (Continued)

miR-281-3p	3.316493205	0.008964708	0.985315433	1	0.118701798	0.845442804	1	0.158645041	0.791005686	1
miR-282-3p	8.367271885	-0.05502631	0.894686928	1	-0.083638774	0.872350229	1	0.083824674	0.871631641	1
miR-283-5p	8.889675463	0.184239889	0.661805509	1	-0.178207501	0.729173429	1	0.115359253	0.820774353	1
miR-284-5p	8.528188577	-0.015733189	0.969804434	1	-0.243480051	0.637135486	1	-0.26469791	0.61066807	1
miR-285-3p	6.82271699	-0.033106621	0.938341073	1	-0.219027771	0.688476053	1	-0.042247156	0.938092165	1
miR-285-5p	10.23618039	0.123303267	0.764392858	1	-0.150619114	0.765838275	1	-0.347129471	0.495340538	1
miR-286-3p	3.919218741	-0.228904075	0.637080464	1	0.059603695	0.91999009	1	-0.121531342	0.837530498	1
miR-2a-1-5p	2.329380427	-0.581042572	0.28083852	1	-0.316037206	0.635322752	1	-0.40795922	0.537260123	1
miR-2a-2-5p	6.423892038	-1.095063684	0.017884934	0.323916019	-0.824663421	0.16846089	1	-0.731096367	0.232604204	1
miR-2a-3p	6.452557627	-0.100257453	0.815826051	1	0.208321228	0.698449039	1	-0.112623103	0.834475809	1
miR-2b-1-5p	1.370915036	0.002847814	0.997763456	1	0.439018007	0.578323868	1	-0.128611601	0.860916459	1
miR-2b-2-5p	6.097843173	-0.316651553	0.466181054	1	-0.174823161	0.748632041	1	-0.125540783	0.818690101	1
miR-2b-3p	7.264307529	0.017241204	0.967262864	1	0.182830116	0.733758532	1	-0.082354966	0.877558807	1
miR-2c-5p	4.818247587	0.131526549	0.776875125	1	0.356948899	0.527478242	1	-0.212473558	0.709919844	1
miR-304-5p	8.186291826	-0.154930554	0.707267538	1	0.566696537	0.278513944	1	0.322057596	0.540468035	1
miR-305-3p	8.218779781	-0.033589442	0.937399666	1	-0.349772227	0.525145132	1	-0.259848235	0.633843596	1
miR-305-5p	8.051862412	-0.290734661	0.487140561	1	-0.359473678	0.498409205	1	-0.159672038	0.761050581	1
miR-306-3p	5.881937572	-0.102994572	0.818094379	1	-0.167856199	0.763742444	1	-0.705557653	0.216918191	1
miR-306-5p	6.109294119	-0.835918592	0.069514035	0.871599056	-0.748232547	0.226180934	1	-0.468082292	0.454563584	1
miR-307a-3p	8.499935706	-0.318882684	0.44523638	1	-0.593352574	0.255089079	1	-0.329651318	0.524007087	1
miR-307a-5p	2.263030857	0.469968189	0.372834328	1	0.230172697	0.740051303	1	0.687452684	0.319573382	1
miR-307b-3p	1.830528351	-0.587992357	0.314938562	1	-0.690178655	0.365127322	1	-0.612664914	0.420215664	1
miR-308-3p	3.191713274	-0.875474543	0.089739447	1	-0.003400586	0.994166613	1	0.661577176	0.320055735	1
miR-312-3p	3.945523735	-1.659199692	0.001289438	0.042035678	-0.705716964	0.297386889	1	-1.318339372	0.045947594	0.832161977
miR-314-3p	8.37376838	1.443989577	0.000982691	0.040044656	0.548143742	0.280283683	1	1.477616407	0.003407092	0.185118642
miR-314-5p	2.284241089	1.686189628	0.005674363	0.132131584	0.548037252	0.405630076	1	1.578457236	0.015208886	0.413174731
miR-315-5p	11.46591991	-0.188529457	0.641974107	1	-0.315794262	0.526760409	1	-0.401067305	0.422406374	1
miR-316-3p	5.21356517	-0.152611281	0.750916013	1	-0.443841244	0.504859243	1	-0.618420609	0.373689784	1
miR-316-5p	8.090571659	0.09974866	0.810757067	1	0.077509849	0.883950597	1	0.189112141	0.719009104	1
miR-317-3p	13.17322299	-0.434995885	0.311484714	1	-1.075173331	0.05791588	1	-0.857250889	0.135853947	1
miR-317-5p	5.230829214	-0.150027263	0.737378599	1	-0.154415484	0.785272651	1	-0.313975724	0.576284665	1
miR-318-3p	3.496263345	-1.190305361	0.022320504	0.363824212	-1.21221768	0.062351367	1	-1.347656284	0.035754045	0.74328897
miR-31a-3p	1.6540048	0.101555761	0.857034134	1	-0.141050444	0.85187446	1	-0.054749733	0.940869503	1
miR-31a-5p	9.9035515	-0.071160717	0.861400609	1	0.246780269	0.659215579	1	-0.099309168	0.850914252	1
miR-31b-5p	4.793109398	0.178603871	0.705975355	1	-0.260080858	0.657233017	1	-0.520813916	0.381931423	1
miR-33-3p	4.668544478	-0.047694361	0.917444372	1	-0.008445212	0.988497279	1	0.253130576	0.65924018	1
miR-33-5p	11.48048597	0.05256508	0.895185518	1	-0.08285812	0.868027311	1	0.047676531	0.923706952	1
miR-34-3p	8.063707758	-0.104140247	0.806250248	1	-0.207895554	0.695308962	1	-0.565805583	0.289735058	1
miR-34-5p	12.20270936	0.376297206	0.359275638	1	-0.234880729	0.646506739	1	-0.313659967	0.545020103	1
miR-375-3p	8.005550386	0.214387882	0.619259215	1	-0.724247684	0.209343945	1	-0.6534049	0.260780784	1
miR-375-5p	5.265727028	0.380188878	0.410860053	1	-0.280214706	0.617798277	1	-0.580253266	0.304522318	1
miR-4919-5p	1.862810986	0.168397763	0.763050042	1	-0.728671933	0.335225358	1	-1.586494309	0.03648044	0.74328897
miR-4951-5p	6.216678219	0.070815681	0.873726611	1	-0.143796137	0.792033071	1	-0.403891691	0.460980347	1
miR-4960-3p	5.131359312	0.118807574	0.795096608	1	-0.131525624	0.816419107	1	-0.384866944	0.496809071	1
miR-4969-5p	2.364800835	0.340384369	0.506821976	1	0.330983656	0.608654349	1	-0.146377937	0.821819107	1
miR-5-5p	3.36636367	-0.229698736	0.644200273	1	-0.111890738	0.854888576	1	0.01336056	0.983057878	1
miR-7-5p	10.36805848	-0.044280935	0.917068692	1	-0.403319929	0.492821493	1	-0.583803902	0.318043982	1
miR-79-3p	5.455666796	-0.222775942	0.615898992	1	-0.049693811	0.931805589	1	0.17160729	0.780927756	1
miR-79-5p	2.823556623	-0.183766978	0.70733866	1	0.232569174	0.735856949	1	0.075290707	0.90821792	1
miR-8-3p	15.85693044	-0.096992441	0.800548007	1	0.381605323	0.42679733	1	0.274381481	0.567788025	1
miR-8-5p	13.81169041	0.013230751	0.97354938	1	-0.002452417	0.996016196	1	-0.061808502	0.900152934	1
miR-87-3p	4.105457922	-0.160188718	0.759978441	1	-1.088673872	0.140997619	1	-0.695569573	0.334715081	1
miR-87-5p	3.683581286	0.440664482	0.361164037	1	-0.480123123	0.434929452	1	-0.578005341	0.348691219	1
miR-927-3p	10.60948793	-0.269885417	0.517139014	1	-0.254002513	0.616275612	1	-0.451526766	0.375351969	1
miR-927-5p	11.45401412	-0.237436034	0.553127092	1	-0.303940521	0.539705989	1	-0.381314553	0.444711467	1
miR-929-3p	2.506970451	-0.590105603	0.251006266	1	-0.67609194	0.312840293	1	-0.270645622	0.684263945	1
miR-929-5p	9.005498783	-0.17753719	0.661417255	1	0.289837706	0.580221178	1	0.379299505	0.473262554	1
miR-92a-3p	3.27549448	-0.162052517	0.743178332	1	0.345295557	0.57585119	1	-0.60495265	0.342049836	1
miR-92a-5p	5.775516775	0.372508293	0.419393215	1	-0.282099037	0.627574084	1	-0.474261792	0.417965972	1
miR-92b-3p	6.107794862	0.062031414	0.889147878	1	0.134795338	0.80591297	1	-0.504836084	0.366048219	1
miR-932-3p	4.19443899	0.319402631	0.48490241	1	0.171131805	0.782803292	1	-0.189261317	0.754317639	1
miR-932-5p	11.37089801	-0.325020622	0.419990024	1	-0.249216316	0.615171276	1	-0.1895236	0.701503187	1
miR-954-5p	4.365763248	0.116402063	0.803130427	1	-0.303481373	0.604343773	1	-0.649948914	0.270657713	1
miR-956-3p	10.91184837	1.812438154	0.000154324	0.008384943	0.394817835	0.413741093	1	1.490435064	0.001760028	0.143442265

Figure S2.2 (Continued)

miR-956-5p	2.948260573	1.496591064	0.009788804	0.199446882	0.550971407	0.382629954	1	1.589471753	0.010868428	0.357331616
miR-957-3p	14.89873089	-7.37E-05	0.999872694	1	-0.00478068	0.992153768	1	-0.215858731	0.656442273	1
miR-957-5p	2.85383034	-0.035817123	0.942489985	1	-0.223167393	0.719452201	1	-0.844868224	0.180799386	1
miR-958-3p	8.310831727	2.314706243	4.92E-07	8.01E-05	0.186303652	0.712606082	1	1.265770139	0.010961093	0.357331616
miR-958-5p	6.900025359	2.158785348	4.46E-06	0.000363861	1.090557475	0.038676418	1	1.897431458	0.000267576	0.043614936
miR-965-3p	2.185727415	0.115132026	0.825147595	1	-0.648687471	0.360035701	1	0.531833202	0.418877148	1
miR-965-5p	8.673057205	-0.040078405	0.923648695	1	-0.130946098	0.802189004	1	-0.428793312	0.413671224	1
miR-966-5p	3.492204108	0.261042147	0.588434242	1	0.03542036	0.953668679	1	-0.404828471	0.517074282	1
miR-967-5p	3.800271284	-0.596168265	0.21974432	1	-0.2194901	0.737515533	1	-0.506737358	0.428234025	1
miR-969-3p	3.67426177	-0.054542991	0.9093359	1	-0.404694795	0.492689429	1	-0.547509734	0.353481786	1
miR-970-3p	9.889888449	0.363498277	0.378513642	1	-0.052915808	0.918301256	1	-0.148200385	0.774953424	1
miR-970-5p	2.989430389	0.354780971	0.465227858	1	0.190790354	0.768211605	1	-0.232657342	0.724766606	1
miR-971-3p	2.745983975	0.032224011	0.948608821	1	0.967916908	0.124149682	1	-0.136368314	0.835278486	1
miR-971-5p	4.255314796	0.091513821	0.845332306	1	0.316918761	0.593395279	1	-0.236249476	0.692707371	1
miR-980-5p	6.239027074	-0.126544708	0.78115459	1	0.098283546	0.860483976	1	-0.277972542	0.620261724	1
miR-981-3p	8.204984709	-0.548648316	0.195904533	1	-0.460475873	0.417696922	1	-0.536414522	0.336176169	1
miR-981-5p	2.327424776	-0.027767015	0.958064638	1	-0.56018062	0.409068799	1	0.139995321	0.829152245	1
miR-982-5p	2.798511974	0.513856202	0.320590307	1	0.517863506	0.422471861	1	0.282794795	0.657643077	1
miR-983-5p	4.459624063	-0.192373735	0.682554385	1	0.014548354	0.980527413	1	-0.03235354	0.955003657	1
miR-987-5p	12.77368322	0.103532919	0.796941135	1	-0.32950657	0.509275876	1	-0.623660981	0.212891317	1
miR-988-3p	9.120133801	-0.192739887	0.646834886	1	-0.324100961	0.537902496	1	-0.485078023	0.357074917	1
miR-988-5p	6.638418198	-0.140393635	0.742488919	1	-0.475359725	0.387008862	1	-0.462366792	0.405683546	1
miR-989-3p	1.982007788	-1.614402128	0.005450829	0.132131584	-0.209989858	0.759678602	1	-0.647151365	0.350071959	1
miR-990-5p	5.12536801	-0.142470623	0.752513182	1	-0.36325654	0.519076221	1	-0.557394543	0.323625118	1
miR-993-3p	4.95980935	-0.158083802	0.724474946	1	0.336653703	0.561078486	1	-0.639712541	0.279184378	1
miR-993-5p	4.189138895	0.641609359	0.166784246	1	0.255821688	0.679395815	1	-0.043012814	0.945277263	1
miR-995-3p	6.371078462	0.345076459	0.431002717	1	0.084222541	0.87863987	1	0.019490568	0.971788725	1
miR-995-5p	1.394820354	-0.110990302	0.847077846	1	0.172152181	0.812713162	1	-0.488236208	0.508668712	1
miR-996-3p	8.912290788	-0.253338497	0.537764884	1	-0.081567862	0.873915698	1	0.110503492	0.82909609	1
miR-996-5p	4.013570169	-0.063634969	0.894674482	1	-0.091170523	0.880868253	1	-0.506443425	0.412489136	1
miR-998-3p	6.479387215	0.170441745	0.694320491	1	0.017025782	0.975188479	1	-0.150266435	0.783190351	1
miR-998-5p	5.112213837	-0.092990825	0.838908985	1	-0.727846552	0.228106048	1	-0.649981224	0.280682624	1
miR-999-3p	12.96543783	-0.650871594	0.10503707	1	-0.317882125	0.527620197	1	-0.294609246	0.558875817	1
miR-9a-5p	12.1277482	-0.017140672	0.965272802	1	0.03524345	0.943533191	1	-0.100000626	0.843449747	1
miR-9b-3p	6.591093721	-0.122764621	0.780341338	1	-0.543443031	0.32666275	1	-0.62827422	0.253683262	1
miR-9b-5p	9.44835821	0.020327894	0.960646576	1	-0.313673167	0.54172067	1	-0.28295317	0.580542927	1
miR-9c-5p	11.62107167	-0.008822764	0.982556318	0.999872694	-0.400528039	0.42379021	1	-0.404696698	0.420383735	1

Differential expression analysis for canonical mature microRNAs was conducted using the edgeR Bioconductor software package.(*2*) 5 canonical mature microRNAs are expressed at levels that differ significantly from control in any condition.

Figure S2.3 DIANA MicroT target predictions for regulated miRNAs

FlyBase ID	Symbol	dme-miR-312-3p	dme-miR-314-3p	dme-miR-956-3p	dme-miR-958-3p	dme-miR-958-5p
FBgn0000157	Dll	1	0	0	0	0
FBgn0003382	sha	1	0	0	0	0
FBgn0004396	CrebA	1	0	0	0	0
FBgn0038439	Cad89D	1	0	0	0	0
FBgn0032001	CG8360	1	0	0	0	0
FBgn0034143	CG8303	1	0	0	0	0
FBgn0019968	Khc-73	1	0	0	0	0
FBgn0036741	CG7510	1	0	0	0	0
FBgn0039810	CG15549	1	0	0	0	0
FBgn0033867	Cpr50Ca	1	0	0	0	0
FBgn0038438	Der-2	1	0	0	0	0
FBgn0029835	CG5921	1	0	0	0	0
FBgn0051122	CG31122	1	0	1	0	0
FBgn0085227	CG34198	1	0	0	0	0
FBgn0001319	kn	1	0	1	0	0
FBgn0038073	CG14395	1	0	0	0	0
FBgn0035283	CG12024	1	0	0	1	0
FBgn0035625	Blimp-1	1	0	0	0	0
FBgn0051224	CG31224	1	0	0	0	0
FBgn0017549	0	1	1	0	0	0
FBgn0033996	CG11807	1	0	0	0	0
FBgn0028642	#N/A	1	0	0	1	0
FBgn0039530	Tusp	1	0	0	0	0
FBgn0086475	sec3	1	0	0	0	0
FBgn0035050	ST6Gal	1	0	0	0	0
FBgn0015567	#N/A	1	0	0	1	0
FBgn0259789	vfl	1	0	0	0	0
FBgn0035144	CG17181	1	1	0	0	0
FBgn0033903	CG8323	1	0	0	0	0
FBgn0034997	CG3376	1	0	0	0	0
FBgn0036935	CG14186	1	0	0	0	0
FBgn0037672	sage	1	0	0	0	0
FBgn0015772	Nak	1	0	0	1	0
FBgn0085250	CG34221	1	0	0	0	0
FBgn0038418	pad	1	0	0	0	0
FBgn0030921	CG6290	1	0	0	0	0
FBgn0040359	CG11380	1	0	0	0	0
FBgn0038815	CG5466	1	0	0	0	0
FBgn0037744	CG8417	1	0	0	0	0
FBgn0034509	Obp57c	1	0	0	0	0
FBgn0037416	Osi9	1	0	0	0	0
FBgn0262636	Lin29	1	0	1	0	0
FBgn0038463	CG3534	1	0	0	0	0
FBgn0260451	CG14042	1	0	0	0	0
FBgn0004656	fs(1)h	1	0	0	0	0
FBgn0039808	CG12071	1	0	0	0	0
FBgn0051632	sens-2	1	0	0	0	0
FBgn0023441	fus	1	0	0	0	0
FBgn0051637	CG31637	1	1	0	0	0
FBgn0031457	CG3077	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0031375	erm	1	0	0	0	0
FBgn0010350	CdsA	1	0	0	0	0
FBgn0261239	Hr39	1	0	0	0	0
FBgn0040532	CG8369	1	0	0	0	0
FBgn0039431	CG6490	1	0	1	0	0
FBgn0016794	dos	1	0	0	0	0
FBgn0003870	ttk	1	0	1	0	0
FBgn0030668	CG8128	1	0	0	0	0
FBgn0008654	0	1	0	1	0	0
FBgn0262737	mub	1	0	1	0	0
FBgn0033609	fbl6	1	0	0	0	0
FBgn0038611	CG14309	1	0	0	0	0
FBgn0005633	fln	1	0	0	0	0
FBgn0003733	tor	1	0	0	0	0
FBgn0011656	Mef2	1	0	0	0	0
FBgn0029846	#N/A	1	0	0	0	0
FBgn0004101	bs	1	0	0	0	0
FBgn0034808	CG9896	1	0	0	0	0
FBgn0032023	CG14274	1	0	0	0	0
FBgn0030408	CG11085	1	0	0	0	0
FBgn0030362	regucalcin	1	0	0	0	0
FBgn0036819	dysb	1	0	0	0	0
FBgn0036323	CG14118	1	0	0	0	0
FBgn0261648	salm	1	0	0	0	0
FBgn0036285	toe	1	0	0	0	0
FBgn0034304	CG5742	1	0	0	0	0
FBgn0261274	Ero1L	1	0	0	0	0
FBgn0038461	CG3678	1	0	0	0	0
FBgn0033448	hebe	1	0	0	0	0
FBgn0259750	#N/A	1	0	0	0	0
FBgn0052105	CG32105	1	0	0	0	0
FBgn0000562	egl	1	0	0	0	0
FBgn0040099	lectin-28C	1	0	0	0	0
FBgn0260963	#N/A	1	0	0	0	0
FBgn0086372	lap	1	0	0	1	0
FBgn0051038	CG31038	1	0	0	0	0
FBgn0004397	Vinc	1	0	0	0	0
FBgn0050463	CG30463	1	0	0	0	0
FBgn0026313	X11L	1	0	0	0	0
FBgn0003209	raw	1	0	0	0	0
FBgn0032513	CG6565	1	0	0	0	0
FBgn0032129	jp	1	1	0	0	0
FBgn0037988	CG14740	1	0	0	0	0
FBgn0015014	#N/A	1	0	0	0	0
FBgn0021895	ytr	1	0	0	0	0
FBgn0039380	CG5890	1	0	0	0	0
FBgn0003162	Pu	1	0	0	0	0
FBgn0035308	CG15822	1	0	0	0	0
FBgn0031698	Ncoa6	1	0	0	0	0
FBgn0028421	Kap3	1	1	0	0	0
FBgn0262740	Evi5	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0086910	0	1	0	0	0	0
FBgn0026616	alpha-Man-IIb	1	0	0	0	0
FBgn0034420	CG10737	1	0	0	0	0
FBgn0004370	Ptp10D	1	0	0	0	0
FBgn0005677	dac	1	0	0	0	0
FBgn0030499	CG11178	1	0	0	0	0
FBgn0259986	nab	1	0	0	0	0
FBgn0004381	Klp68D	1	0	0	0	0
FBgn0035338	CG13800	1	0	0	0	0
FBgn0036360	CG10713	1	0	0	0	0
FBgn0003507	srp	1	1	0	1	0
FBgn0015323	VAChT	1	0	0	0	0
FBgn0259791	bora	1	0	0	0	0
FBgn0025360	Optix	1	0	0	0	0
FBgn0028997	nmdyn-D7	1	0	0	0	0
FBgn0039187	CG6454	1	0	0	0	0
FBgn0004914	Hnf4	1	0	0	0	0
FBgn0040752	Prosap	1	0	0	0	0
FBgn0020309	crol	1	0	0	0	0
FBgn0013759	CASK	1	0	0	0	0
FBgn0015776	nrv1	1	0	0	0	0
FBgn0085414	dpr12	1	0	0	0	0
FBgn0030997	CG7990	1	0	0	0	0
FBgn0033984	Lap1	1	0	0	0	0
FBgn0043841	vir-1	1	0	0	0	0
FBgn0037241	CG14646	1	0	0	0	0
FBgn0260400	elav	1	0	0	1	0
FBgn0015129	0	1	0	0	0	0
FBgn0051869	CG31869	1	0	0	0	0
FBgn0010303	hep	1	0	0	0	0
FBgn0039257	tnc	1	0	0	0	0
FBgn0034271	CG4996	1	0	0	1	0
FBgn0034327	CG14505	1	0	0	0	0
FBgn0000567	Eip74EF	1	0	0	0	0
FBgn0030778	CG4678	1	0	0	0	0
FBgn0023215	Mnt	1	0	0	0	0
FBgn0030004	CG10958	1	0	0	0	0
FBgn0030774	spheroide	1	0	0	0	0
FBgn0036544	sff	1	0	0	0	0
FBgn0051005	qless	1	0	0	0	0
FBgn0039852	nyo	1	0	1	0	0
FBgn0051191	CG31191	1	0	0	0	0
FBgn0035540	Syx17	1	0	0	0	0
FBgn0024238	Fim	1	0	0	0	0
FBgn0005640	Eip63E	1	0	0	0	0
FBgn0039887	CG2053	1	0	0	0	0
FBgn0052600	dpr8	1	0	0	0	0
FBgn0028872	CG18095	1	0	0	0	0
FBgn0000120	Arr1	1	0	0	0	0
FBgn0052638	CG32638	1	0	0	0	1
FBgn0038787	CG4360	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0263097	Glut4EF	1	0	0	0	0
FBgn0037101	CG7634	1	0	0	0	0
FBgn0039132	AP-1sigma	1	0	0	0	0
FBgn0029893	CG14442	1	0	0	0	0
FBgn0261986	RASSF8	1	1	1	0	0
FBgn0034057	CG8314	1	0	0	0	0
FBgn0262350	#N/A	1	1	0	0	0
FBgn0004198	ct	1	0	0	0	1
FBgn0013948	#N/A	1	0	0	0	0
FBgn0052553	CG32553	1	0	0	0	0
FBgn0024555	flfl	1	0	0	0	0
FBgn0033138	Tsp42Eq	1	0	0	0	0
FBgn0038281	RpL10Aa	1	0	0	0	0
FBgn0262573	#N/A	1	0	0	0	0
FBgn0262582	cic	1	0	0	0	0
FBgn0039528	dsd	1	0	0	0	0
FBgn0036623	CG4729	1	0	0	0	0
FBgn0052119	CG32119	1	0	0	0	0
FBgn0043070	MESK2	1	0	0	0	0
FBgn0014870	Psi	1	0	0	0	1
FBgn0002283	l(3)73Ah	1	0	0	0	0
FBgn0003254	rib	1	0	0	0	0
FBgn0034433	endoB	1	0	0	1	0
FBgn0004873	#N/A	1	0	0	0	0
FBgn0050392	CG30392	1	0	0	0	0
FBgn0087007	bbg	1	0	0	0	0
FBgn0003060	CG9757	1	0	0	0	0
FBgn0032340	Ge-1	1	0	0	0	0
FBgn0003149	Prm	1	0	0	0	0
FBgn0259231	CCKLR-17D1	1	1	0	0	0
FBgn0028500	Rich	1	0	0	0	0
FBgn0024944	Oamb	1	0	0	0	0
FBgn0020245	0	1	0	0	0	0
FBgn0036927	CG7433	1	0	0	0	0
FBgn0028996	onecut	1	0	0	0	0
FBgn0015229	glec	1	0	0	0	0
FBgn0052195	CG32195	1	0	0	0	0
FBgn0259937	Nop60B	1	0	0	0	0
FBgn0004569	aos	1	0	0	0	0
FBgn0037107	CG7166	1	0	0	0	0
FBgn0034230	CG4853	1	0	0	0	0
FBgn0038658	CG14292	1	0	0	0	0
FBgn0036958	CG17233	1	0	0	0	0
FBgn0036271	Pbgs	1	0	0	0	0
FBgn0262353	CR43051	1	0	0	0	0
FBgn0029761	SK	1	1	1	0	1
FBgn0031453	Bacc	1	0	0	0	0
FBgn0000064	Ald	1	0	0	0	0
FBgn0035574	RhoGEF64C	1	0	0	0	0
FBgn0053543	CG33543	1	0	0	0	0
FBgn0038296	CG6752	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0044823	Spec2	1	0	0	0	0
FBgn0050424	CG30424	1	0	0	0	0
FBgn0086677	jeb	1	0	0	0	0
FBgn0036483	CG12316	1	0	0	0	0
FBgn0085424	nub	1	0	0	0	0
FBgn0004242	Syt1	1	1	0	0	0
FBgn0039911	CG1909	1	0	0	0	0
FBgn0029996	Ubc-E2H	1	0	0	0	0
FBgn0085376	CG34347	1	0	0	0	0
FBgn0003715	0	1	0	0	1	0
FBgn0050147	Hil	1	0	0	0	0
FBgn0038332	CG6136	1	0	0	0	0
FBgn0000037	mAcR	1	0	0	0	0
FBgn0261836	Msp-300	1	0	0	0	0
FBgn0037428	Osi18	1	0	0	0	0
FBgn0085404	CG34375	1	0	0	0	0
FBgn0040395	Unc-76	1	0	0	0	0
FBgn0036770	Prestin	1	0	0	0	0
FBgn0025821	l-t	1	0	0	0	0
FBgn0028541	TM9SF4	1	0	0	0	0
FBgn0036814	CG14073	1	0	0	0	0
FBgn0031304	CG4552	1	0	0	0	0
FBgn0015778	rin	1	0	0	0	0
FBgn0262729	#N/A	1	0	0	0	0
FBgn0083949	CG34113	1	0	0	0	0
FBgn0039078	CG4374	1	1	1	0	0
FBgn0035016	CG4612	1	0	0	0	0
FBgn0000210	br	1	0	0	0	0
FBgn0039958	CG12567	1	0	0	0	0
FBgn0000541	E(bx)	1	0	0	0	0
FBgn0000317	ck	1	0	0	0	0
FBgn0086736	GckIII	1	0	0	0	0
FBgn0259935	CG42458	1	0	0	0	0
FBgn0261560	Thor	1	0	0	0	0
FBgn0013983	imd	1	0	0	0	0
FBgn0250910	Octbeta3R	1	0	0	0	0
FBgn0025186	ari-2	1	0	0	0	0
FBgn0034399	CG15083	1	0	0	0	0
FBgn0085383	CG34354	1	0	0	0	0
FBgn0040697	Teh3	1	0	0	0	0
FBgn0051536	#N/A	1	0	0	0	0
FBgn0035270	CG13933	1	0	0	0	0
FBgn0038388	CG4287	1	0	0	0	0
FBgn0033589	CG13227	1	0	0	0	0
FBgn0032700	CG10338	1	0	0	1	0
FBgn0000547	ed	1	0	0	0	0
FBgn0010905	Spn	1	0	0	1	0
FBgn0030617	CG9095	1	0	0	0	0
FBgn0000394	cv	1	0	0	0	0
FBgn0015623	Cpr	1	0	0	0	0
FBgn0050080	CG30080	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0032021	CG7781	1	0	0	1	0
FBgn0015919	caup	1	0	0	0	0
FBgn0038109	CG11656	1	0	0	0	0
FBgn0000524	dx	1	0	0	1	0
FBgn0086378	Alg-2	1	0	0	0	0
FBgn0041004	CG17715	1	0	0	0	0
FBgn0031981	CG7466	1	0	0	0	0
FBgn0029715	CG11444	1	0	0	0	0
FBgn0035429	CG12017	1	0	0	0	0
FBgn0040334	Tsp3A	1	0	0	0	0
FBgn0044826	Pak3	1	0	0	0	1
FBgn0261244	inaE	1	0	0	0	0
FBgn0035229	CG7852	1	1	0	1	0
FBgn0029861	CG3815	1	0	0	0	0
FBgn0038124	CG14380	1	0	0	1	0
FBgn0039704	neo	1	1	0	0	0
FBgn0039411	dys	1	0	0	0	0
FBgn0038818	Nep4	1	0	0	0	0
FBgn0012051	CalpA	1	0	0	0	0
FBgn0020389	Papss	1	0	0	0	0
FBgn0022935	D19A	1	0	0	0	0
FBgn0005671	Vha55	1	0	0	0	0
FBgn0023407	B4	1	0	0	0	0
FBgn0039161	CG13606	1	0	0	0	0
FBgn0261786	mi	1	0	0	0	0
FBgn0001138	grn	1	0	0	0	0
FBgn0051140	CG31140	1	0	1	0	0
FBgn0263117	CG34377	1	0	1	0	0
FBgn0031676	CG14040	1	0	0	0	0
FBgn0036760	CG5567	1	0	0	0	0
FBgn0033906	CG8331	1	0	0	0	0
FBgn0030252	#N/A	1	0	0	0	0
FBgn0030532	#N/A	1	0	0	0	0
FBgn0039634	alpha-Man-Ib	1	0	0	1	0
FBgn0000577	en	1	0	0	0	0
FBgn0025724	beta'Cop	1	0	0	0	0
FBgn0030788	Sap30	1	0	0	0	0
FBgn0030680	CG8944	1	1	1	0	0
FBgn0003334	Scm	1	0	0	0	0
FBgn0004197	Ser	1	0	0	0	0
FBgn0035490	CG1136	1	0	0	0	0
FBgn0038304	CG12241	1	0	0	0	0
FBgn0038339	CG6118	1	0	0	0	0
FBgn0046704	Liprin-alpha	1	0	0	0	0
FBgn0028875	nAcRalpha-34E	1	0	0	0	0
FBgn0033677	CG8321	1	0	0	0	0
FBgn0031478	CG8814	1	0	0	0	0
FBgn0001259	in	1	0	0	0	0
FBgn0041203	LIMK1	1	0	0	0	0
FBgn0052445	CG32445	1	0	0	0	0
FBgn0039730	CG7903	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0015278	Pi3K68D	1	0	0	0	0
FBgn0030654	CG15643	1	0	0	0	0
FBgn0040816	CG12521	1	0	0	0	0
FBgn0033608	CG13220	1	0	0	0	0
FBgn0000286	Cf2	1	0	0	0	0
FBgn0029092	ced-6	1	0	0	0	1
FBgn0010280	Taf4	1	0	0	0	0
FBgn0028371	jbug	1	0	0	0	0
FBgn0035772	Sh3beta	1	0	0	0	0
FBgn0029970	Nek2	1	0	0	0	0
FBgn0037847	SelR	1	0	0	0	0
FBgn0040271	Sulf1	1	0	0	0	0
FBgn0041210	HDAC4	1	0	0	0	0
FBgn0030745	CG4239	1	0	0	0	0
FBgn0030756	CG9903	1	0	0	0	0
FBgn0035253	CG7971	1	0	0	0	0
FBgn0067864	Patj	1	0	0	0	0
FBgn0028999	nerfin-1	1	0	1	0	0
FBgn0004436	UbcD6	1	0	0	0	0
FBgn0034926	#N/A	1	0	0	0	0
FBgn0015295	shark	1	0	0	0	0
FBgn0003301	rut	1	0	0	0	0
FBgn0030930	GalNAc-T2	1	0	0	0	0
FBgn0002921	Atpalpha	1	0	0	0	0
FBgn0001324	kto	1	0	0	0	0
FBgn0067628	CG33331	1	0	0	0	0
FBgn0040230	dbo	1	0	0	0	1
FBgn0038420	CG10311	1	0	0	0	0
FBgn0038435	Gyc-89Da	1	0	0	0	0
FBgn0030812	CG8949	1	0	0	0	0
FBgn0037279	CG1129	1	0	0	0	0
FBgn0031760	Tsp26A	1	0	0	1	0
FBgn0003721	Tm1	1	0	0	0	0
FBgn0043536	Obp57d	1	0	0	0	0
FBgn0040502	CG8343	1	0	0	0	0
FBgn0033958	CG12858	1	0	0	0	0
FBgn0037993	dpr15	1	0	0	0	0
FBgn0032140	CG13117	1	0	0	0	0
FBgn0027835	Dp1	1	0	0	0	0
FBgn0032601	yellow-b	1	0	0	0	0
FBgn0040343	CG3713	1	0	0	0	0
FBgn0259749	mmy	1	0	0	0	0
FBgn0028704	Nckx30C	1	1	0	0	0
FBgn0032341	Reps	1	0	1	1	0
FBgn0259708	CG42362	1	0	0	0	0
FBgn0085371	CG34342	1	0	0	0	0
FBgn0050446	Tdc2	1	0	0	0	0
FBgn0033058	CCHa2r	1	0	0	0	0
FBgn0001297	kay	1	0	0	0	0
FBgn0000054	Adf1	1	0	0	0	0
FBgn0085196	CG34167	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0036761	MED19	1	0	0	0	0
FBgn0038460	CG18622	1	0	0	0	0
FBgn0033465	Etf-QO	1	0	0	0	0
FBgn0262716	Arp3	1	0	0	0	0
FBgn0260466	Indy-2	1	0	0	0	0
FBgn0260470	SP555	1	0	0	0	0
FBgn0036008	CG3408	1	0	0	0	0
FBgn0037794	CG6254	1	0	0	0	0
FBgn0003042	Pc	1	0	0	0	0
FBgn0026386	Or47a	1	0	0	0	0
FBgn0004598	Fur2	1	0	0	0	0
FBgn0053203	CG33203	1	0	0	0	0
FBgn0085432	pan	1	0	0	0	0
FBgn0030447	CG2200	1	0	0	0	0
FBgn0053196	dp	1	0	0	0	0
FBgn0052150	CG32150	1	0	0	0	0
FBgn0033222	CG12824	1	0	0	0	0
FBgn0086472	RpS25	1	0	0	0	0
FBgn0028582	lqf	1	0	0	0	0
FBgn0260941	app	1	0	0	0	0
FBgn0031636	CG12194	1	0	0	0	0
FBgn0003028	ovo	1	1	0	0	0
FBgn0030296	CG15196	1	0	0	0	0
FBgn0036334	CG11267	1	0	0	0	0
FBgn0011742	Arp2	1	0	0	0	0
FBgn0045852	ham	1	0	0	0	0
FBgn0051342	CG31342	1	0	0	0	0
FBgn0010381	Drs	1	0	0	0	0
FBgn0036488	CG6878	1	0	0	0	0
FBgn0042083	CG3267	1	0	0	0	0
FBgn0034517	Cpr57A	1	0	0	0	0
FBgn0030018	slpr	1	0	0	0	0
FBgn0011300	babo	1	0	0	0	0
FBgn0030913	CG6123	1	1	0	0	0
FBgn0005619	Hdc	1	0	0	0	0
FBgn0037326	CG14669	1	0	0	0	0
FBgn0014343	mirr	1	0	0	0	0
FBgn0040208	Kat60	1	0	0	0	0
FBgn0052702	CG32702	1	0	0	0	0
FBgn0031882	Rab30	1	0	0	0	0
FBgn0086901	cv-c	1	0	0	0	0
FBgn0000117	arm	1	0	0	0	0
FBgn0034957	CG3121	1	0	0	0	0
FBgn0031395	CG10874	1	0	0	0	0
FBgn0000119	arr	1	0	0	0	0
FBgn0033382	Hydr1	1	0	0	0	0
FBgn0259211	grh	1	0	0	0	0
FBgn0035383	CG2107	1	0	0	0	0
FBgn0034072	Dg	1	0	0	0	0
FBgn0013733	shot	1	0	0	1	0
FBgn0261552	ps	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0004395	unk	1	0	0	0	0
FBgn0028879	CG15270	1	0	0	0	0
FBgn0000546	EcR	1	0	1	0	0
FBgn0004924	Top1	1	0	0	0	0
FBgn0027885	Aac11	1	0	0	0	0
FBgn0262511	Vha44	1	0	0	0	0
FBgn0036165	chrb	1	0	0	0	0
FBgn0259142	0	1	0	0	0	0
FBgn0020762	Atet	1	0	0	0	0
FBgn0026206	mei-P26	1	0	0	0	0
FBgn0000273	Pka-C1	1	0	0	0	0
FBgn0046247	CG5938	1	0	0	0	0
FBgn0000097	aop	1	0	0	0	0
FBgn0000464	Lar	1	0	0	0	0
FBgn0040759	CG13177	1	0	0	0	0
FBgn0032178	Spn31A	1	0	0	0	0
FBgn0004666	sim	1	0	0	0	0
FBgn0039580	Gfat2	1	0	0	0	0
FBgn0033853	CG6145	1	0	0	0	0
FBgn0000028	acj6	1	0	0	0	0
FBgn0039776	PH4alphaEFB	1	0	0	1	0
FBgn0052651	CG32651	1	0	0	0	0
FBgn0026084	cib	1	0	0	0	0
FBgn0000303	Cha	1	0	0	0	0
FBgn0085434	NaCP60E	1	0	0	0	0
FBgn0003016	osp	1	0	0	0	1
FBgn0031950	Herp	1	0	0	0	0
FBgn0037120	CG11247	1	0	0	0	0
FBgn0004647	N	1	0	0	0	0
FBgn0001291	Jra	1	0	0	0	0
FBgn0028916	CG33090	1	0	0	0	0
FBgn0036688	Fit2	1	0	0	0	0
FBgn0262738	norpA	1	1	1	0	0
FBgn0085385	CG34356	1	1	0	0	0
FBgn0038858	CG5793	1	0	0	0	0
FBgn0032901	sky	1	0	0	0	0
FBgn0035388	CG2162	1	0	0	0	0
FBgn0026371	SAK	1	0	0	0	0
FBgn0052579	CG32579	1	0	0	0	0
FBgn0031879	uif	1	0	0	0	0
FBgn0027535	botv	1	0	0	0	0
FBgn0036485	FucTA	1	0	0	0	0
FBgn0024244	drm	1	0	0	0	0
FBgn0031174	CG1486	1	0	0	0	0
FBgn0261555	CG42673	1	0	0	0	0
FBgn0027594	drpr	1	0	0	0	0
FBgn0015591	Ast	1	0	0	0	0
FBgn0031232	CG11617	1	0	0	0	0
FBgn0262739	AGO1	1	1	0	0	0
FBgn0035987	CG3689	1	0	0	1	0
FBgn0035955	CG5194	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0033876	Syngr	1	0	0	0	0
FBgn0015774	NetB	1	0	0	1	0
FBgn0262111	f	1	0	0	0	0
FBgn0035241	CG12105	1	0	0	0	0
FBgn0025726	unc-13	1	0	0	0	0
FBgn0010473	tutl	1	0	0	0	0
FBgn0086359	Invadolysin	1	0	0	0	0
FBgn0037842	CG6567	1	0	0	0	0
FBgn0030985	Obp18a	1	0	0	0	0
FBgn0031298	Atg4	1	0	0	0	0
FBgn0035170	dpr20	1	0	0	0	0
FBgn0052085	CG32085	1	0	0	0	0
FBgn0037472	CG10098	1	0	0	0	0
FBgn0259739	CG42393	1	0	0	0	0
FBgn0051313	CG31313	1	0	0	0	0
FBgn0001970	Pgant35A	1	0	0	0	0
FBgn0022770	Peritrophin-A	1	0	0	0	0
FBgn0037636	CG9821	1	1	0	0	0
FBgn0030049	Trf4-1	1	0	0	0	0
FBgn0011740	alpha-Man-II	1	0	0	1	0
FBgn0031770	CG13995	1	0	0	0	0
FBgn0034266	CG4975	1	0	0	1	0
FBgn0051915	CG31915	1	0	0	0	0
FBgn0053556	form3	1	0	0	1	0
FBgn0015799	Rbf	1	0	0	0	0
FBgn0261999	CG42817	1	0	0	0	0
FBgn0037622	CG8202	1	0	0	0	0
FBgn0038341	CG14869	1	1	1	0	0
FBgn0037917	wkd	1	0	0	0	0
FBgn0031592	Art2	1	0	0	0	0
FBgn0011826	Pp2B-14D	1	0	0	0	0
FBgn0033915	CG8485	1	0	0	0	0
FBgn0034425	CG11906	1	0	0	0	0
FBgn0036556	CG5830	1	0	0	0	0
FBgn0037848	Tsp86D	1	0	0	0	0
FBgn0052767	CG32767	1	0	0	0	0
FBgn0029508	Tsp42Ea	1	0	0	1	0
FBgn0000253	Cam	1	0	1	0	0
FBgn0062978	CG31808	1	0	0	0	0
FBgn0033638	CG9005	1	0	0	0	0
FBgn0263042	CG43337	1	0	0	0	0
FBgn0250848	26-29-p	1	0	0	0	0
FBgn0037516	CG11286	1	0	0	0	0
FBgn0010417	Taf6	1	0	0	0	0
FBgn0023423	slmb	1	0	0	0	0
FBgn0037364	Rab23	1	0	0	0	0
FBgn0033279	CG2291	1	0	0	0	0
FBgn0050419	CG30419	1	0	0	0	0
FBgn0051949	CG31949	1	0	0	0	0
FBgn0261387	CG17528	1	0	0	0	0
FBgn0032409	Ced-12	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0024277	trio	1	0	0	0	0
FBgn0037130	Syn1	1	0	0	0	0
FBgn0035593	CG4603	1	0	0	0	0
FBgn0028431	#N/A	1	0	0	0	0
FBgn0015513	mbc	1	0	0	0	0
FBgn0002577	m	1	0	0	0	0
FBgn0032120	CG33298	1	0	0	0	1
FBgn0031675	CG9121	1	0	0	0	0
FBgn0031310	CG4764	1	0	0	0	0
FBgn0038237	Pde6	1	0	0	0	0
FBgn0036948	CG7298	1	0	0	0	0
FBgn0038893	CG6353	1	0	0	0	0
FBgn0024362	CG11412	1	0	0	0	0
FBgn0085405	CG34376	1	0	0	0	0
FBgn0038880	SIFR	1	0	0	0	0
FBgn0011582	Dop1R1	1	1	1	1	0
FBgn0000448	Hr46	1	1	0	0	0
FBgn0033607	CG9062	1	0	0	0	0
FBgn0015872	Drip	1	0	0	0	0
FBgn0022268	KdelR	1	0	0	0	0
FBgn0262896	CG43251	1	0	0	0	0
FBgn0039925	Kif3C	1	0	0	0	0
FBgn0031260	Spp	1	0	0	0	0
FBgn0261553	CG42671	1	1	0	0	0
FBgn0028476	CG15817	1	0	0	0	0
FBgn0031257	CG4133	1	0	0	0	1
FBgn0000221	brn	1	0	0	0	0
FBgn0035529	CG1319	1	0	0	0	0
FBgn0029962	CG1402	1	0	0	0	0
FBgn0260812	inaF-D	1	0	0	0	0
FBgn0038652	CG7720	1	0	0	0	0
FBgn0051523	CG31523	1	0	0	0	0
FBgn0032987	RpL21	1	0	0	0	0
FBgn0031811	CG13982	1	0	0	0	0
FBgn0037556	CG9636	1	1	0	0	0
FBgn0011676	Nos	1	0	0	0	0
FBgn0038324	CG5038	1	0	0	0	0
FBgn0023000	mth	1	0	0	0	0
FBgn0026259	eIF5B	1	0	0	0	0
FBgn0032003	CG8349	1	0	0	0	0
FBgn0053516	dpr3	1	0	0	0	0
FBgn0052452	CG32452	1	0	0	0	0
FBgn0011584	Trp1	1	0	0	0	0
FBgn0005278	Sam-S	1	0	0	0	0
FBgn0024814	Clc	1	0	0	0	0
FBgn0032763	CG17568	1	0	0	0	0
FBgn0026086	Adar	1	0	1	0	0
FBgn0033693	CG13175	1	0	0	0	0
FBgn0263131	CG43373	1	0	1	0	0
FBgn0035504	Teh4	1	1	0	0	0
FBgn0023091	dimm	1	0	0	0	0

Figure S2.3 (Continued)

FBgn0052574	Twdlalpha	1	0	0	0	0
FBgn0042630	Sox21b	1	0	0	0	1
FBgn0053555	#N/A	0	1	0	0	0
FBgn0013753	Bgb	0	1	0	0	0
FBgn0024234	gbb	0	1	1	0	0
FBgn0262579	Ect4	0	1	1	0	0
FBgn0040388	boi	0	1	1	0	0
FBgn0023081	gek	0	1	1	0	0
FBgn0032702	CG10376	0	1	0	0	0
FBgn0037336	CG2519	0	1	1	1	0
FBgn0032895	twit	0	1	0	0	0
FBgn0083961	CG34125	0	1	0	0	0
FBgn0000490	dpp	0	1	0	0	0
FBgn0038118	timeout	0	1	0	0	0
FBgn0036610	CG13058	0	1	0	0	0
FBgn0004449	Ten-m	0	1	0	0	0
FBgn0040682	CG14664	0	1	0	0	0
FBgn0263218	0	0	1	1	0	0
FBgn0036402	CG6650	0	1	0	0	0
FBgn0038705	CG11626	0	1	0	0	0
FBgn0026147	CG16833	0	1	0	0	0
FBgn0259210	prom	0	1	0	0	0
FBgn0085450	Snoo	0	1	0	0	0
FBgn0028484	Ack	0	1	0	0	0
FBgn0260499	qvr	0	1	1	0	0
FBgn0037856	CG4674	0	1	0	0	0
FBgn0034158	CG5522	0	1	0	0	0
FBgn0034691	synj	0	1	0	0	0
FBgn0035151	CG17129	0	1	0	0	0
FBgn0031816	CG16947	0	1	0	0	0
FBgn0085400	CG34371	0	1	1	0	0
FBgn0004839	otk	0	1	0	0	0
FBgn0038545	CG7713	0	1	0	0	0
FBgn0027364	Six4	0	1	0	0	0
FBgn0024321	NK7.1	0	1	0	0	0
FBgn0261456	hpo	0	1	0	0	0
FBgn0085397	Fili	0	1	0	0	0
FBgn0260632	dl	0	1	0	0	0
FBgn0015396	jumu	0	1	0	0	0
FBgn0086253	rumi	0	1	0	0	0
FBgn0027546	CG4766	0	1	0	0	0
FBgn0037722	CG8319	0	1	0	0	0
FBgn0030482	CG1673	0	1	0	0	0
FBgn0032374	CG14931	0	1	0	0	0
FBgn0259985	Mppe	0	1	0	0	0
FBgn0038320	Sra-1	0	1	0	0	0
FBgn0029514	312	0	1	0	0	0
FBgn0011211	blw	0	1	0	0	0
FBgn0025632	CG4313	0	1	0	0	0
FBgn0032485	CG9426	0	1	0	0	0
FBgn0004622	Takr99D	0	1	1	0	0

Figure S2.3 (Continued)

FBgn0035510	Cpr64Aa	0	1	0	0	0
FBgn0035148	CG3402	0	1	0	0	0
FBgn0010894	sinu	0	1	0	0	0
FBgn0003071	Pfk	0	1	0	0	1
FBgn0037408	NPFR1	0	1	0	0	0
FBgn0259244	CG42342	0	1	0	0	0
FBgn0035997	phol	0	1	0	0	0
FBgn0085335	#N/A	0	1	0	0	0
FBgn0033901	O-fut1	0	1	0	0	0
FBgn0031149	CG1518	0	1	0	0	0
FBgn0032955	CG2201	0	1	0	0	0
FBgn0035359	CG1143	0	1	0	0	0
FBgn0027512	CG10254	0	1	1	0	0
FBgn0260934	par-1	0	1	0	0	0
FBgn0010548	Aldh-III	0	1	0	0	0
FBgn0029878	Pat1	0	1	0	0	0
FBgn0030309	CG1572	0	1	0	0	0
FBgn0051960	CG31960	0	1	0	0	0
FBgn0004892	sob	0	1	0	1	0
FBgn0050035	Tret1-1	0	1	0	0	0
FBgn0028550	Atf3	0	1	1	0	0
FBgn0044028	Notum	0	1	0	0	0
FBgn0016059	Sema-1b	0	1	1	0	0
FBgn0031632	CG15628	0	1	1	0	0
FBgn0015268	Nap1	0	1	0	0	0
FBgn0036564	#N/A	0	1	0	0	0
FBgn0040777	CG14767	0	1	0	0	0
FBgn0034720	Liprin-gamma	0	1	0	0	0
FBgn0000667	Actn	0	1	0	0	0
FBgn0036317	CG10948	0	1	0	1	0
FBgn0038874	ETHR	0	1	0	0	0
FBgn0263116	5-HT1B	0	1	0	0	0
FBgn0028506	CG4455	0	1	1	0	0
FBgn0051716	Cnot4	0	1	0	0	0
FBgn0032723	ssp3	0	1	0	0	0
FBgn0038829	CG17271	0	1	0	1	0
FBgn0083991	CG34155	0	1	0	0	0
FBgn0039532	Mtl	0	1	0	0	0
FBgn0030877	Arp8	0	1	0	0	0
FBgn0037475	Fer1	0	1	0	0	0
FBgn0033652	ths	0	1	0	0	0
FBgn0033827	CG17047	0	1	0	0	0
FBgn0086783	#N/A	0	1	0	0	0
FBgn0036450	CG13472	0	1	0	0	0
FBgn0250862	CG42237	0	1	0	0	0
FBgn0037643	skap	0	1	0	0	0
FBgn0039945	CG17159	0	1	0	0	0
FBgn0036518	RhoGAP71E	0	1	0	0	0
FBgn0032934	CG8679	0	1	0	0	0
FBgn0030932	Ggt-1	0	1	0	0	0
FBgn0040305	MTF-1	0	1	0	0	0

Figure S2.3 (Continued)

FBgn0035583	CG13704	0	1	0	0	0
FBgn0003206	Ras64B	0	1	0	0	0
FBgn0013988	#N/A	0	1	0	0	0
FBgn0036922	CG14182	0	1	0	0	0
FBgn0262160	CG9932	0	1	0	0	0
FBgn0010389	htl	0	1	0	0	0
FBgn0003459	stwl	0	1	0	0	0
FBgn0034472	CG8517	0	1	0	0	0
FBgn0027951	MTA1-like	0	1	0	0	0
FBgn0038595	CG7142	0	1	1	0	0
FBgn0053558	mim	0	1	1	0	0
FBgn0010435	emp	0	1	0	0	0
FBgn0016792	dmt	0	1	0	0	0
FBgn0010238	Lac	0	1	0	0	0
FBgn0260012	pds5	0	1	0	0	0
FBgn0261804	CG42750	0	1	0	0	0
FBgn0025469	slv	0	1	0	0	0
FBgn0014930	CG2846	0	1	0	0	0
FBgn0052062	A2bp1	0	1	0	0	0
FBgn0035424	CG11505	0	1	0	0	0
FBgn0039152	Rootletin	0	1	0	0	0
FBgn0039283	danr	0	1	0	0	0
FBgn0035914	CG6282	0	1	0	0	0
FBgn0030432	CG4404	0	1	0	0	1
FBgn0051778	CG31778	0	1	0	0	0
FBgn0034135	Syn2	0	1	0	0	0
FBgn0037031	CG11456	0	1	0	0	0
FBgn0033347	CG8248	0	1	0	0	0
FBgn0261859	CG42788	0	1	0	0	0
FBgn0024963	GluClalpha	0	1	0	0	1
FBgn0039249	CG11168	0	1	0	0	0
FBgn0035475	CG10866	0	1	0	0	0
FBgn0030766	mthl1	0	1	0	0	0
FBgn0014135	bnl	0	1	0	0	0
FBgn0050362	boly	0	1	0	0	0
FBgn0040232	cmet	0	1	0	0	0
FBgn0037188	CG7369	0	1	0	0	0
FBgn0029705	CG12693	0	1	0	0	0
FBgn0067102	GlcT-1	0	1	0	0	0
FBgn0028360	l(1)G0148	0	1	0	0	0
FBgn0026430	Grip84	0	1	0	0	0
FBgn0034606	ASPP	0	1	0	0	0
FBgn0263132	Cht6	0	1	0	0	0
FBgn0004366	0	0	1	0	0	0
FBgn0011758	B-H1	0	1	0	0	0
FBgn0028496	CG30116	0	1	0	0	0
FBgn0024732	Drep-1	0	1	0	0	0
FBgn0040465	Dip3	0	1	0	0	0
FBgn0030053	CG12081	0	1	0	0	0
FBgn0039454	CG14247	0	1	0	0	0
FBgn0031995	CG8475	0	1	0	0	0

Figure S2.3 (Continued)

FBgn0037249	eIF3-S10	0	1	0	0	0
FBgn0026136	CkIIbeta2	0	1	0	0	0
FBgn0004648	svr	0	1	0	0	0
FBgn0002940	ninaE	0	1	0	0	0
FBgn0036522	CG7372	0	1	1	0	0
FBgn0036446	CG9384	0	1	0	1	0
FBgn0003210	rb	0	1	0	0	0
FBgn0032797	CG10186	0	1	0	0	0
FBgn0015372	RabX1	0	1	0	0	0
FBgn0011725	twin	0	1	0	0	0
FBgn0022987	qkr54B	0	1	0	0	0
FBgn0035085	CG3770	0	1	0	0	1
FBgn0030961	CG7058	0	1	0	0	0
FBgn0054051	CG34051	0	1	0	0	0
FBgn0261053	Cad86C	0	1	0	0	0
FBgn0039335	Vps33B	0	1	0	0	0
FBgn0031498	CG17260	0	1	0	0	0
FBgn0034049	bdg	0	1	0	0	0
FBgn0016047	nompA	0	1	0	0	0
FBgn0035372	CG12093	0	1	0	0	0
FBgn0032378	CycY	0	1	1	1	0
FBgn0031001	CG7884	0	1	0	0	0
FBgn0039435	TwdlP	0	1	0	0	0
FBgn0031195	CG17600	0	1	0	0	0
FBgn0085442	NKAIN	0	1	0	0	0
FBgn0026869	Thd1	0	1	0	0	0
FBgn0038720	CG6231	0	1	0	0	0
FBgn0035235	CG7879	0	1	0	0	0
FBgn0017418	ari-1	0	1	0	0	0
FBgn0033285	CG18449	0	1	0	0	0
FBgn0039970	CG17508	0	1	0	0	0
FBgn0024754	Flo-1	0	1	0	0	0
FBgn0028683	spt4	0	1	0	0	0
FBgn0039816	CG11317	0	1	0	0	0
FBgn0037680	CG8121	0	1	0	0	0
FBgn0039014	CG6982	0	1	0	0	0
FBgn0022355	Tsf1	0	1	0	0	0
FBgn0032633	Lrch	0	1	0	0	0
FBgn0030485	CG1998	0	1	0	0	0
FBgn0040505	Alk	0	1	0	0	0
FBgn0034025	GalNAc-T1	0	1	0	0	0
FBgn0030864	CG8173	0	1	0	0	0
FBgn0263112	Mitf	0	1	0	0	0
FBgn0039527	CG5639	0	1	0	0	0
FBgn0038139	PK2-R2	0	1	0	0	0
FBgn0030544	CG13403	0	1	0	0	0
FBgn0000015	Abd-B	0	1	0	0	0
FBgn0263111	cac	0	1	0	0	0
FBgn0039229	Saf-B	0	1	0	0	0
FBgn0045202	#N/A	0	1	0	0	0
FBgn0003371	sgg	0	1	0	0	0

Figure S2.3 (Continued)

FBgn0005672	spi	0	1	0	0	0
FBgn0000633	fas	0	1	0	0	0
FBgn0261085	Syt12	0	1	0	0	0
FBgn0038149	CG9796	0	1	0	0	0
FBgn0037440	CG1041	0	1	0	0	0
FBgn0031637	CG2950	0	1	0	1	0
FBgn0003944	Ubx	0	1	0	0	0
FBgn0250850	rig	0	1	0	0	0
FBgn0035101	p130CAS	0	1	0	0	0
FBgn0016641	PTP-ER	0	0	1	0	1
FBgn0003710	tipE	0	0	1	0	0
FBgn0001250	if	0	0	1	0	0
FBgn0086687	desat1	0	0	1	0	0
FBgn0038089	d-cup	0	0	1	0	0
FBgn0035464	CG12006	0	0	1	0	0
FBgn0036260	Rh7	0	0	1	0	0
FBgn0040009	CG17490	0	0	1	0	0
FBgn0261914	#N/A	0	0	1	0	0
FBgn0036353	CG10171	0	0	1	0	0
FBgn0001612	l(1)dd4	0	0	1	0	0
FBgn0035285	CG12025	0	0	1	0	0
FBgn0261245	sing	0	0	1	0	0
FBgn0040089	meso18E	0	0	1	1	0
FBgn0259938	cwo	0	0	1	1	0
FBgn0030357	Sclp	0	0	1	0	0
FBgn0262722	CG43166	0	0	1	0	0
FBgn0002643	mam	0	0	1	0	0
FBgn0025743	mbt	0	0	1	0	0
FBgn0038065	Snx3	0	0	1	0	0
FBgn0031885	Mnn1	0	0	1	0	0
FBgn0038453	CG10326	0	0	1	0	0
FBgn0036202	CG6024	0	0	1	1	0
FBgn0029974	dpr14	0	0	1	0	0
FBgn0036951	CG7017	0	0	1	0	0
FBgn0052177	Ndfip	0	0	1	0	0
FBgn0037614	CG8116	0	0	1	0	0
FBgn0029957	CG12155	0	0	1	0	0
FBgn0033155	Br140	0	0	1	0	0
FBgn0038243	CG8066	0	0	1	0	0
FBgn0261514	NimA	0	0	1	0	0
FBgn0038890	CG7956	0	0	1	1	0
FBgn0053995	CG33995	0	0	1	0	0
FBgn0031948	CG7149	0	0	1	0	0
FBgn0051262	CG31262	0	0	1	0	0
FBgn0260954	CG42586	0	0	1	0	0
FBgn0003165	pum	0	0	1	0	0
FBgn0028537	CG31775	0	0	1	0	0
FBgn0039157	Myo95E	0	0	1	0	0
FBgn0262718	#N/A	0	0	1	0	0
FBgn0020299	stumps	0	0	1	0	0
FBgn0032025	CG7778	0	0	1	0	0

Figure S2.3 (Continued)

FBgn0031434	insv	0	0	1	0	0
FBgn0034435	CG9975	0	0	1	0	0
FBgn0000635	Fas2	0	0	1	0	0
FBgn0031397	CG15385	0	0	1	0	0
FBgn0031674	#N/A	0	0	1	0	0
FBgn0052815	CG32815	0	0	1	0	0
FBgn0053202	dpr11	0	0	1	0	0
FBgn0053207	pxb	0	0	1	0	0
FBgn0031119	CG1812	0	0	1	0	0
FBgn0032281	CG17107	0	0	1	0	0
FBgn0023416	Ac3	0	0	1	0	0
FBgn0036762	CG7430	0	0	1	0	0
FBgn0259225	#N/A	0	0	1	0	0
FBgn0032476	CG5439	0	0	1	0	0
FBgn0030346	CG11802	0	0	1	0	0
FBgn0031097	obst-A	0	0	1	0	0
FBgn0034546	CG13442	0	0	1	0	0
FBgn0263102	psq	0	0	1	0	0
FBgn0013432	bcn92	0	0	1	0	0
FBgn0085446	CG34417	0	0	1	1	0
FBgn0259209	Mlp60A	0	0	1	0	0
FBgn0000363	#N/A	0	0	1	0	0
FBgn0035245	GC	0	0	1	0	0
FBgn0034286	dpr13	0	0	1	0	0
FBgn0033426	CG1814	0	0	1	0	0
FBgn0259234	Camta	0	0	1	0	0
FBgn0034763	RYBP	0	0	1	0	0
FBgn0020513	ade5	0	0	1	0	0
FBgn0011225	jar	0	0	1	0	0
FBgn0004456	mew	0	0	1	0	0
FBgn0085470	lmgB	0	0	1	0	0
FBgn0262734	Rbp2	0	0	1	0	1
FBgn0032297	CG17124	0	0	1	0	0
FBgn0261477	slim	0	0	1	0	0
FBgn0262742	Fas1	0	0	1	0	0
FBgn0000395	cv-2	0	0	1	0	1
FBgn0039069	CG6763	0	0	1	0	0
FBgn0262871	lute	0	0	1	0	0
FBgn0014340	mof	0	0	1	0	0
FBgn0034262	swi2	0	0	1	0	0
FBgn0027508	Tnks	0	0	1	0	0
FBgn0033524	Cyp49a1	0	0	1	0	0
FBgn0004228	mex1	0	0	1	0	0
FBgn0045770	S-Lap3	0	0	1	0	0
FBgn0039927	CG11155	0	0	1	0	0
FBgn0053056	CG33056	0	0	1	0	0
FBgn0024941	RSG7	0	0	1	0	0
FBgn0035998	CG3437	0	0	1	0	0
FBgn0010313	corto	0	0	1	0	0
FBgn0015402	ksr	0	0	1	0	0
FBgn0259699	CG42353	0	0	1	0	0

Figure S2.3 (Continued)

FBgn0034978	CG3257	0	0	1	0	0
FBgn0086779	step	0	0	1	0	0
FBgn0001247	Ide	0	0	1	0	0
FBgn0037138	P5CDh1	0	0	0	1	0
FBgn0001169	H	0	0	0	1	0
FBgn0003093	Pkc98E	0	0	0	1	0
FBgn0032587	CG5953	0	0	0	1	0
FBgn0259100	#N/A	0	0	0	1	0
FBgn0020278	loco	0	0	0	1	0
FBgn0022764	Sin3A	0	0	0	1	0
FBgn0039113	CG10217	0	0	0	1	0
FBgn0262473	Tl	0	0	0	1	0
FBgn0026189	prominin-like	0	0	0	1	0
FBgn0020907	Scp2	0	0	0	1	0
FBgn0003861	trp	0	0	0	1	0
FBgn0043364	cbt	0	0	0	1	0
FBgn0038295	Gyc88E	0	0	0	1	0
FBgn0046874	Pif1B	0	0	0	1	0
FBgn0015371	chn	0	0	0	1	0
FBgn0050280	CG30280	0	0	0	1	0
FBgn0040636	CG13255	0	0	0	1	0
FBgn0042180	CG18870	0	0	0	1	0
FBgn0036677	CG13023	0	0	0	1	0
FBgn0037659	Kdm2	0	0	0	1	1
FBgn0028407	Drep-3	0	0	0	1	0
FBgn0035078	Tpc2	0	0	0	1	0
FBgn0034267	CG4984	0	0	0	1	0
FBgn0020236	ATPCL	0	0	0	1	0
FBgn0053156	CG33156	0	0	0	1	0
FBgn0034300	CG5098	0	0	0	1	0
FBgn0035145	MED14	0	0	0	1	0
FBgn0015777	nrv2	0	0	0	1	0
FBgn0004607	zfh2	0	0	0	1	0
FBgn0033321	CG8738	0	0	0	1	0
FBgn0086346	ALiX	0	0	0	1	0
FBgn0261673	nemy	0	0	0	1	0
FBgn0043799	CG31381	0	0	0	1	0
FBgn0004875	enc	0	0	0	1	0
FBgn0026320	Tom	0	0	0	1	0
FBgn0002931	net	0	0	0	1	0
FBgn0086365	Orct2	0	0	0	1	0
FBgn0030758	CanA-14F	0	0	0	1	1
FBgn0003041	pbl	0	0	0	1	0
FBgn0250851	CG33981	0	0	0	1	0
FBgn0000588	esc	0	0	0	1	0
FBgn0040281	Aplip1	0	0	0	1	0
FBgn0040319	Gclc	0	0	0	1	0
FBgn0036273	CG10426	0	0	0	1	0
FBgn0032796	CG10188	0	0	0	1	0
FBgn0001941	ifc	0	0	0	1	0
FBgn0032312	CG14071	0	0	0	1	1

Figure S2.3 (Continued)

FBgn0034583	CG10527	0	0	0	1	0
FBgn0262614	pyd	0	0	0	1	0
FBgn0261015	Pif1A	0	0	0	1	0
FBgn0050423	CG30423	0	0	0	1	0
FBgn0034491	Hsl	0	0	0	1	0
FBgn0011746	ana	0	0	0	1	0
FBgn0028406	Drep-4	0	0	0	1	0
FBgn0020930	Dgkepsilon	0	0	0	1	0
FBgn0037235	CG1103	0	0	0	1	0
FBgn0030033	CG1387	0	0	0	1	0
FBgn0085313	CG34284	0	0	0	1	0
FBgn0030327	FucT6	0	0	0	1	0
FBgn0013305	Nmda1	0	0	0	1	0
FBgn0037521	CG2993	0	0	0	1	0
FBgn0033679	CG8888	0	0	0	1	0
FBgn0027570	Nep2	0	0	0	1	0
FBgn0033391	CG8026	0	0	0	1	0
FBgn0036732	Oatp74D	0	0	0	1	1
FBgn0010453	Wnt4	0	0	0	1	0
FBgn0033686	Hen1	0	0	0	1	0
FBgn0017397	#N/A	0	0	0	1	0
FBgn0024320	Npc1a	0	0	0	1	0
FBgn0039644	CG11897	0	0	0	1	0
FBgn0037015	cmpy	0	0	0	1	0
FBgn0037552	CG7800	0	0	0	1	0
FBgn0015793	Rab19	0	0	0	1	0
FBgn0039596	CG10000	0	0	0	1	0
FBgn0039632	Cul-5	0	0	0	1	0
FBgn0004611	Plc21C	0	0	0	1	0
FBgn0033627	CG13204	0	0	0	1	0
FBgn0086676	spin	0	0	0	1	0
FBgn0040079	pkaap	0	0	0	1	0
FBgn0035239	CG18170	0	0	0	1	0
FBgn0086708	stv	0	0	0	1	0
FBgn0000414	Dab	0	0	0	1	0
FBgn0039467	CG14253	0	0	0	1	0
FBgn0030590	CG9518	0	0	0	1	0
FBgn0261625	CG42708	0	0	0	1	0
FBgn0033569	CG12942	0	0	0	1	0
FBgn0036789	AlCR2	0	0	0	1	0
FBgn0051217	modSP	0	0	0	1	0
FBgn0036290	CG10638	0	0	0	1	0
FBgn0033483	egr	0	0	0	1	0
FBgn0032514	CG9302	0	0	0	1	0
FBgn0039056	CenB1A	0	0	0	1	0
FBgn0000346	comt	0	0	0	1	0
FBgn0029729	CG12682	0	0	0	1	0
FBgn0038088	CG10126	0	0	0	1	0
FBgn0011829	Ret	0	0	0	1	0
FBgn0000542	ec	0	0	0	1	0
FBgn0038826	Syp	0	0	0	1	0

Figure S2.3 (Continued)

FBgn0028509	CenG1A	0	0	0	1	0
FBgn0033856	CG13334	0	0	0	1	0
FBgn0261526	NT1	0	0	0	1	0
FBgn0032123	Oatp30B	0	0	0	1	0
FBgn0038721	CG16718	0	0	0	1	0
FBgn0038693	unc79	0	0	0	1	0
FBgn0017551	Rca1	0	0	0	1	0
FBgn0028525	c(2)M	0	0	0	1	0
FBgn0011764	Dsp1	0	0	0	1	0
FBgn0031857	CG11321	0	0	0	1	0
FBgn0035678	CG10469	0	0	0	1	0
FBgn0001087	g	0	0	0	1	0
FBgn0000636	Fas3	0	0	0	1	0
FBgn0020240	#N/A	0	0	0	1	0
FBgn0041147	ida	0	0	0	1	0
FBgn0038619	CG7685	0	0	0	1	0
FBgn0034027	CG8187	0	0	0	1	0
FBgn0039831	CG12054	0	0	0	1	0
FBgn0259822	Ca-beta	0	0	0	1	0
FBgn0051760	CG31760	0	0	0	1	0
FBgn0016754	sba	0	0	0	1	0
FBgn0033473	CG12128	0	0	0	1	0
FBgn0050118	CG30118	0	0	0	1	0
FBgn0003638	su(w[a])	0	0	0	1	0
FBgn0011259	Sema-1a	0	0	0	1	0
FBgn0014859	Hr38	0	0	0	1	0
FBgn0025682	scf	0	0	0	1	0
FBgn0053129	CG33129	0	0	0	1	0
FBgn0015609	CadN	0	0	0	1	0
FBgn0032200	CG5676	0	0	0	1	0
FBgn0028433	Ggamma30A	0	0	0	1	0
FBgn0032817	CG10631	0	0	0	1	0
FBgn0003732	Top2	0	0	0	1	0
FBgn0034731	CG10384	0	0	0	1	0
FBgn0033375	CG8078	0	0	0	1	0
FBgn0035157	CG13894	0	0	0	1	0
FBgn0040376	CG13759	0	0	0	1	0
FBgn0036843	CG6812	0	0	0	1	0
FBgn0004373	fwd	0	0	0	1	0
FBgn0041096	rols	0	0	0	1	0
FBgn0011837	Tis11	0	0	0	1	0
FBgn0034570	CG10543	0	0	0	1	0
FBgn0010399	Nmdar1	0	0	0	1	0
FBgn0003513	ss	0	0	0	1	0
FBgn0035830	CG8209	0	0	0	1	0
FBgn0030912	CG6023	0	0	0	1	0
FBgn0261090	Sytbeta	0	0	0	1	0
FBgn0033739	Dyb	0	0	0	1	0
FBgn0028400	Syt4	0	0	0	0	1
FBgn0011818	oaf	0	0	0	0	1
FBgn0004587	B52	0	0	0	0	1

Figure S2.3 (Continued)

FBgn0004865	Eip78C	0	0	0	0	1
FBgn0037796	CG12814	0	0	0	0	1
FBgn0261602	RpL8	0	0	0	0	1
FBgn0038391	GATAe	0	0	0	0	1
FBgn0038498	beat-IIa	0	0	0	0	1
FBgn0034447	CG7744	0	0	0	0	1
FBgn0012034	AcCoAS	0	0	0	0	1
FBgn0029762	NAAT1	0	0	0	0	1
FBgn0029807	CG3108	0	0	0	0	1
FBgn0020238	14-3-3epsilon	0	0	0	0	1
FBgn0085384	CG34355	0	0	0	0	1
FBgn0260430	CG42525	0	0	0	0	1
FBgn0001253	ImpE1	0	0	0	0	1
FBgn0261837	pre-mod(mdg4)-T	0	0	0	0	1
FBgn0030503	Tango2	0	0	0	0	1
FBgn0019990	Gcn2	0	0	0	0	1
FBgn0015831	Rtnl2	0	0	0	0	1
FBgn0085436	Not1	0	0	0	0	1
FBgn0039851	mey	0	0	0	0	1
FBgn0031724	CG18266	0	0	0	0	1
FBgn0034307	CG10914	0	0	0	0	1
FBgn0013325	RpL11	0	0	0	0	1
FBgn0263197	Patronin	0	0	0	0	1
FBgn0085429	#N/A	0	0	0	0	1
FBgn0041605	cpx	0	0	0	0	1
FBgn0036259	CG9760	0	0	0	0	1
FBgn0030328	Amun	0	0	0	0	1
FBgn0085248	CG34219	0	0	0	0	1
FBgn0035056	spz6	0	0	0	0	1
FBgn0034795	MED23	0	0	0	0	1
FBgn0033597	Cpr47Ea	0	0	0	0	1
FBgn0030367	Cyp311a1	0	0	0	0	1
FBgn0036940	obst-J	0	0	0	0	1
FBgn0011760	ctp	0	0	0	0	1
FBgn0016976	stnA	0	0	0	0	1
FBgn0261838	pre-mod(mdg4)-Z	0	0	0	0	1
FBgn0025631	moody	0	0	0	0	1
FBgn0263255	#N/A	0	0	0	0	1
FBgn0051882	CG31882	0	0	0	0	1
FBgn0027836	Dgp-1	0	0	0	0	1
FBgn0036381	CG8745	0	0	0	0	1
FBgn0013984	InR	0	0	0	0	1
FBgn0038225	soti	0	0	0	0	1
FBgn0020258	ppk	0	0	0	0	1
FBgn0087035	AGO2	0	0	0	0	1
FBgn0001108	Gl	0	0	0	0	1
FBgn0262975	cnc	0	0	0	0	1
FBgn0039538	CG12883	0	0	0	0	1
FBgn0085387	shakB	0	0	0	0	1
FBgn0011666	msi	0	0	0	0	1
FBgn0037414	Osi7	0	0	0	0	1

Figure S2.3 (Continued)

FBgn0052369	CG32369	0	0	0	0	1
FBgn0039024	CG4721	0	0	0	0	1
FBgn0016975	stnB	0	0	0	0	1
FBgn0036786	skl	0	0	0	0	1
FBgn0027490	D12	0	0	0	0	1
FBgn0033497	CG12912	0	0	0	0	1
FBgn0039697	CG7834	0	0	0	0	1
FBgn0003651	svp	0	0	0	0	1
FBgn0036566	ClC-c	0	0	0	0	1
FBgn0040087	p115	0	0	0	0	1
FBgn0032286	CG7300	0	0	0	0	1
FBgn0259203	CG42307	0	0	0	0	1
FBgn0035246	CG13928	0	0	0	0	1
FBgn0033088	PGAP3	0	0	0	0	1
FBgn0259246	brp	0	0	0	0	1
FBgn0037135	CG7414	0	0	0	0	1
FBgn0037853	CG14696	0	0	0	0	1
FBgn0027932	Akap200	0	0	0	0	1
FBgn0020372	TM4SF	0	0	0	0	1
FBgn0026058	OdsH	0	0	0	0	1

Target gene predictions were obtained using DIANA micro-T cds for all differentially expressed canonical mature microRNAs.(*3*)

Figure S2.4 PATHER pathway analysis for genes targeted by regulated miRNAs

Pathway	Drosophila REFLIST	Target Genes	Target Genes (expected)	over/under	FDR Adj P-value
Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway	42	12	3.38	+	3.67E-02
Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway	23	8	1.85	+	1.18E-01
Inflammation mediated by chemokine and cytokine signaling pathway	59	12	4.75	+	6.40E-01
Thyrotropin-releasing hormone receptor signaling pathway	24	7	1.93	+	6.63E-01
Wnt signaling pathway	126	20	10.15	+	6.88E-01
Huntington disease	95	16	7.65	+	9.52E-01
Alzheimer disease-amyloid secretase pathway	32	4	2.58	+	1.00E+00
Alpha adrenergic receptor signaling pathway	8	2	0.64	+	1.00E+00
Adrenaline and noradrenaline biosynthesis	17	2	1.37	+	1.00E+00
Nicotine pharmacodynamics pathway	20	3	1.61	+	1.00E+00
Toll_pathway_drosophila	27	3	2.18	+	1.00E+00
SCW_signaling_pathway	21	1	1.69	-	1.00E+00
MYO_signaling_pathway	10	1	0.81	+	1.00E+00
GBB_signaling_pathway	13	1	1.05	-	1.00E+00
DPP_signaling_pathway	22	2	1.77	+	1.00E+00
DPP-SCW_signaling_pathway	24	2	1.93	+	1.00E+00
BMP_signaling_pathway-drosophila	30	4	2.42	+	1.00E+00
Xanthine and guanine salvage pathway	3	0	0.24	-	1.00E+00
Activinbetasignaling_pathway	11	1	0.89	+	1.00E+00
Vitamin B6 metabolism	5	0	0.4	-	1.00E+00
Valine biosynthesis	1	1	0.08	+	1.00E+00
Tyrosine biosynthesis	3	0	0.24	-	1.00E+00
Tryptophan biosynthesis	1	0	0.08	-	1.00E+00
Triacylglycerol metabolism	1	1	0.08	+	1.00E+00
Thiamine metabolism	3	1	0.24	+	1.00E+00
ALP23B_signaling_pathway	10	1	0.81	+	1.00E+00
Synaptic_vesicle_trafficking	12	4	0.97	+	1.00E+00
GABA-B_receptor_II_signaling	9	0	0.73	-	1.00E+00
Endogenous_cannabinoid_signaling	1	0	0.08	-	1.00E+00
Sulfate assimilation	1	1	0.08	+	1.00E+00
Succinate to proprionate conversion	1	1	0.08	+	1.00E+00
Serine glycine biosynthesis	4	0	0.32	-	1.00E+00
Salvage pyrimidine ribonucleotides	16	3	1.29	+	1.00E+00
Salvage pyrimidine deoxyribonucleotides	5	2	0.4	+	1.00E+00
S adenosyl methionine biosynthesis	1	1	0.08	+	1.00E+00
Pyruvate metabolism	25	3	2.01	+	1.00E+00
Pyrimidine Metabolism	14	2	1.13	+	1.00E+00
Pyridoxal phosphate salvage pathway	4	0	0.32	-	1.00E+00
Purine metabolism	13	0	1.05	-	1.00E+00
Proline biosynthesis	2	0	0.16	-	1.00E+00
Phenylethylamine degradation	3	0	0.24	-	1.00E+00
Phenylalanine biosynthesis	2	0	0.16	-	1.00E+00
Pentose phosphate pathway	7	0	0.56	-	1.00E+00
p38 MAPK pathway	12	2	0.97	+	1.00E+00
Opioid proopiomelanocortin pathway	10	0	0.81	-	1.00E+00
Opioid prodynorphin pathway	10	0	0.81	-	1.00E+00
Opioid proenkephalin pathway	9	0	0.73	-	1.00E+00
Nicotine degradation	36	0	2.9	-	1.00E+00
Enkephalin release	14	1	1.13	-	1.00E+00
Dopamine receptor mediated signaling pathway	34	5	2.74	+	1.00E+00
Angiotensin II-stimulated signaling through G proteins and beta-arrestin	5	0	0.4	-	1.00E+00
PLP biosynthesis	4	0	0.32	-	1.00E+00
Ornithine degradation	3	0	0.24	-	1.00E+00
O-antigen biosynthesis	2	1	0.16	+	1.00E+00
N-acetylglucosamine metabolism	4	1	0.32	+	1.00E+00
Methylmalonyl pathway	1	1	0.08	+	1.00E+00
Methylcitrate cycle	4	0	0.32	-	1.00E+00
Methionine biosynthesis	3	0	0.24	-	1.00E+00
Mannose metabolism	5	2	0.4	+	1.00E+00
Lipoate_biosynthesis	5	0	0.4	-	1.00E+00
Ubiquitin proteasome pathway	65	2	5.24	-	1.00E+00
Leucine biosynthesis	1	1	0.08	+	1.00E+00
Isoleucine biosynthesis	2	1	0.16	+	1.00E+00
Heme biosynthesis	14	1	1.13	-	1.00E+00
Glutamine glutamate conversion	4	0	0.32	-	1.00E+00
Fructose galactose metabolism	13	2	1.05	+	1.00E+00
Formyltetrahydroformate biosynthesis	9	0	0.73	-	1.00E+00
Folate biosynthesis	6	1	0.48	+	1.00E+00
Flavin biosynthesis	2	1	0.16	+	1.00E+00
De novo pyrmidine ribonucleotides biosythesis	9	1	0.73	+	1.00E+00
p53 pathway	41	6	3.3	+	1.00E+00
mRNA splicing	4	0	0.32	-	1.00E+00
VEGF signaling pathway	32	4	2.58	+	1.00E+00
Transcription regulation by bZIP transcription factor	43	4	3.46	+	1.00E+00
Toll receptor signaling pathway	17	2	1.37	+	1.00E+00
T cell activation	21	5	1.69	+	1.00E+00
p53 pathway feedback loops 2	27	4	2.18	+	1.00E+00
TGF-beta signaling pathway	44	7	3.54	+	1.00E+00
p53 pathway by glucose deprivation	11	0	0.89	-	1.00E+00
TCA cycle	22	1	1.77	-	1.00E+00
Vitamin D metabolism and pathway	5	2	0.4	+	1.00E+00
Vasopressin synthesis	9	0	0.73	-	1.00E+00
Ras Pathway	31	6	2.5	+	1.00E+00
De novo pyrimidine deoxyribonucleotide biosynthesis	11	1	0.89	+	1.00E+00

Figure S2.4 (Continued)

P53 pathway feedback loops 1	1	1	0.08 +	1.00E+00
De novo purine biosynthesis	20	1	1.61 -	1.00E+00
Oxytocin receptor mediated signaling pathway	21	3	1.69 +	1.00E+00
Cysteine biosynthesis	1	0	0.08 -	1.00E+00
Coenzyme A biosynthesis	4	0	0.32 -	1.00E+00
Carnitine metabolism	2	0	0.16 -	1.00E+00
Carnitine and CoA metabolism	2	0	0.16 -	1.00E+00
Asparagine and aspartate biosynthesis	4	0	0.32 -	1.00E+00
Parkinson disease	80	4	6.44 -	1.00E+00
PI3 kinase pathway	21	4	1.69 +	1.00E+00
PDGF signaling pathway	56	8	4.51 +	1.00E+00
Oxidative stress response	16	0	1.29 -	1.00E+00
Notch signaling pathway	27	2	2.18 -	1.00E+00
Nicotinic acetylcholine receptor signaling pathway	46	8	3.71 +	1.00E+00
Muscarinic acetylcholine receptor 2 and 4 signaling pathway	25	3	2.01 +	1.00E+00
Muscarinic acetylcholine receptor 1 and 3 signaling pathway	29	5	2.34 +	1.00E+00
Histamine synthesis	5	2	0.4 +	1.00E+00
Metabotropic glutamate receptor group I pathway	13	3	1.05 +	1.00E+00
Histamine H2 receptor mediated signaling pathway	12	1	0.97 +	1.00E+00
Metabotropic glutamate receptor group II pathway	21	3	1.69 +	1.00E+00
Histamine H1 receptor mediated signaling pathway	16	3	1.29 +	1.00E+00
Gamma-aminobutyric acid synthesis	5	1	0.4 +	1.00E+00
Ascorbate degradation	6	1	0.48 +	1.00E+00
Arginine biosynthesis	4	0	0.32 -	1.00E+00
Androgen/estrogene/progesterone biosynthesis	30	0	2.42 -	1.00E+00
Cortocotropin releasing factor receptor signaling pathway	12	0	0.97 -	1.00E+00
Aminobutyrate degradation	2	1	0.16 +	1.00E+00
Allantoin degradation	4	0	0.32 -	1.00E+00
Alanine biosynthesis	1	1	0.08 +	1.00E+00
Adenine and hypoxanthine salvage pathway	8	0	0.64 -	1.00E+00
Acetate utilization	2	1	0.16 +	1.00E+00
ATP synthesis	8	1	0.64 +	1.00E+00
Metabotropic glutamate receptor group III pathway	34	5	2.74 +	1.00E+00
Ionotropic glutamate receptor pathway	30	7	2.42 +	1.00E+00
Interleukin signaling pathway	16	4	1.29 +	1.00E+00
Interferon-gamma signaling pathway	11	0	0.89 -	1.00E+00
Integrin signalling pathway	71	11	5.72 +	1.00E+00
Beta3 adrenergic receptor signaling pathway	10	0	0.81 -	1.00E+00
Insulin/IGF pathway-protein kinase B signaling cascade	16	3	1.29 +	1.00E+00
Gonadotropin releasing hormone receptor pathway	69	11	5.56 +	1.00E+00
Beta2 adrenergic receptor signaling pathway	19	1	1.53 -	1.00E+00
Insulin/IGF pathway-mitogen activated protein kinase kinase/MAP kinase cascade	15	3	1.21 +	1.00E+00
Beta1 adrenergic receptor signaling pathway	19	1	1.53 -	1.00E+00
5HT4 type receptor mediated signaling pathway	12	0	0.97 -	1.00E+00
Hypoxia response via HIF activation	20	2	1.61 +	1.00E+00
5HT3 type receptor mediated signaling pathway	7	0	0.56 -	1.00E+00
5HT2 type receptor mediated signaling pathway	29	4	2.34 +	1.00E+00
5HT1 type receptor mediated signaling pathway	18	1	1.45 -	1.00E+00
5-Hydroxytryptamine degredation	6	2	0.48 +	1.00E+00
5-Hydroxytryptamine biosynthesis	8	2	0.64 +	1.00E+00
Heterotrimeric G-protein signaling pathway-rod outer segment phototransduction	16	5	1.29 +	1.00E+00
Hedgehog signaling pathway	22	2	1.77 +	1.00E+00
Glycolysis	23	2	1.85 +	1.00E+00
General transcription regulation	32	3	2.58 +	1.00E+00
General transcription by RNA polymerase I	12	0	0.97 -	1.00E+00
FGF signaling pathway	59	7	4.75 +	1.00E+00
FAS signaling pathway	14	2	1.13 +	1.00E+00
Endothelin signaling pathway	62	9	4.99 +	1.00E+00
EGF receptor signaling pathway	63	9	5.08 +	1.00E+00
DNA replication	20	2	1.61 +	1.00E+00
Cytoskeletal regulation by Rho GTPase	44	4	3.54 +	1.00E+00
Circadian clock system	6	0	0.48 -	1.00E+00
Cholesterol biosynthesis	8	1	0.64 +	1.00E+00
Cell cycle	18	0	1.45 -	1.00E+00
Cadherin signaling pathway	51	7	4.11 +	1.00E+00
Blood coagulation	1	0	0.08 -	1.00E+00
B cell activation	12	4	0.97 +	1.00E+00
Axon guidance mediated by netrin	11	2	0.89 +	1.00E+00
Axon guidance mediated by Slit/Robo	10	2	0.81 +	1.00E+00
Axon guidance mediated by semaphorins	7	2	0.56 +	1.00E+00
Apoptosis signaling pathway	47	7	3.79 +	1.00E+00
Angiogenesis	63	10	5.08 +	1.00E+00
Alzheimer disease-presenilin pathway	40	7	3.22 +	1.00E+00

The PANTHER gene ontology annotation database was used to identify pathways overrepresented in the set of genes targeted by differentially expressed microRNAs.(*4*) This analysis identifies the "Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway" as the only pathway overrepresented.

Figure S2.5 PATHER GO analysis for genes targeted by regulated miRNAs

Molecular Function	Drosophila REFLIST	Target Genes	Target Genes (expected)	over/under	FDR Adj P-value
transcription factor activity	956	140	77.01	+	1.62E-09
transcription regulator activity	956	140	77.01	+	1.62E-09
DNA binding	1114	153	89.74	+	1.63E-08
binding	3360	354	270.67	+	1.45E-06
receptor activity	774	110	62.35	+	1.59E-06
protein binding	1369	161	110.28	+	1.42E-04
structural constituent of cytoskeleton	485	71	39.07	+	2.64E-04
small GTPase regulator activity	197	37	15.87	+	5.08E-04
Unclassified	6361	439	512.42	-	8.07E-04
voltage-gated sodium channel activity	12	7	0.97	+	9.93E-03
kinase activity	355	51	28.6	+	1.18E-02
cytoskeletal protein binding	148	27	11.92	+	1.62E-02
oxidoreductase activity	714	32	57.52	-	1.92E-02
actin binding	96	20	7.73	+	2.30E-02
guanyl-nucleotide exchange factor activity	67	16	5.4	+	2.31E-02
voltage-gated calcium channel activity	15	7	1.21	+	3.85E-02
protein kinase activity	230	35	18.53	+	5.44E-02
ligand-dependent nuclear receptor activity	22	8	1.77	+	7.46E-02
nucleic acid binding	2127	209	171.34	+	1.94E-01
helicase activity	99	1	7.98	-	4.50E-01
motor activity	68	13	5.48	+	6.27E-01
enzyme regulator activity	565	64	45.51	+	7.03E-01
RNA helicase activity	66	0	5.32	-	7.22E-01
receptor binding	443	52	35.69	+	7.98E-01
cation transmembrane transporter activity	285	29	22.96	+	1.00E+00
G-protein coupled receptor activity	241	25	19.41	+	1.00E+00
glucosidase activity	11	2	0.89	+	1.00E+00
transaminase activity	17	4	1.37	+	1.00E+00
galactosidase activity	5	0	0.4	-	1.00E+00
lipid binding	3	0	0.24	-	1.00E+00
DNA ligase activity	4	0	0.32	-	1.00E+00
cytokine receptor activity	7	2	0.56	+	1.00E+00
extracellular matrix structural constituent	29	4	2.34	+	1.00E+00
exoribonuclease activity	25	2	2.01	-	1.00E+00
ATPase activity, coupled to transmembrane movement of substances	86	4	6.93	-	1.00E+00
tumor necrosis factor receptor binding	1	1	0.08	+	1.00E+00
growth factor activity	44	6	3.54	+	1.00E+00
phosphoric diester hydrolase activity	26	4	2.09	+	1.00E+00
transferase activity	1021	99	82.25	+	1.00E+00
exodeoxyribonuclease activity	12	0	0.97	-	1.00E+00
phosphoprotein phosphatase activity	103	15	8.3	+	1.00E+00
endoribonuclease activity	46	1	3.71	-	1.00E+00
endodeoxyribonuclease activity	33	3	2.66	+	1.00E+00
DNA strand annealing activity	5	0	0.4	-	1.00E+00
ligand-gated ion channel activity	94	7	7.57	-	1.00E+00
calcium-dependent phospholipid binding	60	3	4.83	-	1.00E+00
non-membrane spanning protein tyrosine kinase activity	51	7	4.11	+	1.00E+00
transmembrane receptor protein tyrosine kinase activity	35	7	2.82	+	1.00E+00
hydrogen ion transmembrane transporter activity	53	4	4.27	-	1.00E+00
nuclease activity	172	15	13.86	+	1.00E+00
glutamate receptor activity	51	3	4.11	-	1.00E+00
transmembrane receptor protein serine/threonine kinase activity	23	2	1.85	+	1.00E+00
lipase activity	112	4	9.02	-	1.00E+00
acetylcholine receptor activity	24	2	1.93	+	1.00E+00
cysteine-type endopeptidase inhibitor activity	5	2	0.4	+	1.00E+00
serine-type endopeptidase inhibitor activity	62	4	4.99	-	1.00E+00
GABA receptor activity	24	2	1.93	+	1.00E+00
hydrogen ion transporting ATP synthase activity, rotational mechanism	18	3	1.45	+	1.00E+00
RNA splicing factor activity, transesterification mechanism	176	18	14.18	+	1.00E+00
DNA polymerase processivity factor activity	3	0	0.24	-	1.00E+00
ligase activity	367	22	29.56	-	1.00E+00
enzyme inhibitor activity	185	11	14.9	-	1.00E+00
transmembrane transporter activity	696	56	56.07	-	1.00E+00
deaminase activity	26	3	2.09	+	1.00E+00
SNAP receptor activity	22	1	1.77	-	1.00E+00
centromeric DNA binding	14	2	1.13	+	1.00E+00
enzyme activator activity	72	5	5.8	-	1.00E+00
cytokine activity	12	2	0.97	+	1.00E+00

Figure S2.5 (Continued)

catalytic activity	3980	302	320.62 -	1.00E+00
transmembrane receptor protein kinase activity	49	8	3.95 +	1.00E+00
calmodulin binding	155	15	12.49 +	1.00E+00
intramolecular transferase activity	8	0	0.64 -	1.00E+00
lipid transporter activity	30	5	2.42 +	1.00E+00
metallopeptidase activity	181	14	14.58 -	1.00E+00
serine-type peptidase activity	349	16	28.11 -	1.00E+00
cysteine-type peptidase activity	59	4	4.75 -	1.00E+00
peptidase activity	627	34	50.51 -	1.00E+00
ubiquitin-protein ligase activity	139	9	11.2 -	1.00E+00
phosphorylase activity	9	0	0.73 -	1.00E+00
pyrophosphatase activity	11	0	0.89 -	1.00E+00
microtubule motor activity	47	5	3.79 +	1.00E+00
calcium ion binding	243	30	19.58 +	1.00E+00
racemase and epimerase activity	75	1	6.04 -	1.00E+00
isomerase activity	164	6	13.21 -	1.00E+00
intermediate filament binding	1	0	0.08 -	1.00E+00
deacetylase activity	20	2	1.61 +	1.00E+00
phosphatase inhibitor activity	14	2	1.13 +	1.00E+00
phosphatase activator activity	3	0	0.24 -	1.00E+00
kinase inhibitor activity	63	1	5.08 -	1.00E+00
cation channel activity	57	11	4.59 +	1.00E+00
acyltransferase activity	152	12	12.24 -	1.00E+00
kinase activator activity	45	4	3.63 +	1.00E+00
aspartic-type endopeptidase activity	16	1	1.29 -	1.00E+00
phosphatase regulator activity	35	2	2.82 -	1.00E+00
kinase regulator activity	156	10	12.57 -	1.00E+00
microtubule binding	51	7	4.11 +	1.00E+00
nucleotide phosphatase activity	39	1	3.14 -	1.00E+00
carbohydrate phosphatase activity	6	0	0.48 -	1.00E+00
amino acid kinase activity	6	0	0.48 -	1.00E+00
nucleotide kinase activity	48	5	3.87 +	1.00E+00
phospholipase activity	22	3	1.77 +	1.00E+00
RNA methyltransferase activity	20	0	1.61 -	1.00E+00
carbohydrate kinase activity	41	2	3.3 -	1.00E+00
anion channel activity	37	4	2.98 +	1.00E+00
protein disulfide isomerase activity	28	1	2.26 -	1.00E+00
guanylate cyclase activity	26	5	2.09 +	1.00E+00
hydro-lyase activity	47	2	3.79 -	1.00E+00
carboxy-lyase activity	31	3	2.5 +	1.00E+00
hydrolase activity, hydrolyzing N-glycosyl compounds	49	3	3.95 -	1.00E+00
aminoacyl-tRNA ligase activity	36	0	2.9 -	1.00E+00
voltage-gated potassium channel activity	37	1	2.98 -	1.00E+00
methyltransferase activity	104	3	8.38 -	1.00E+00
phosphatase activity	158	16	12.73 +	1.00E+00
amino acid transmembrane transporter activity	106	6	8.54 -	1.00E+00
voltage-gated ion channel activity	60	11	4.83 +	1.00E+00
gap junction channel activity	8	1	0.64 +	1.00E+00
translation release factor activity	6	0	0.48 -	1.00E+00
translation elongation factor activity	64	2	5.16 -	1.00E+00
translation initiation factor activity	74	8	5.96 +	1.00E+00
lyase activity	162	17	13.05 +	1.00E+00
adenylate cyclase activity	41	8	3.3 +	1.00E+00
hydrolase activity, acting on ester bonds	483	41	38.91 +	1.00E+00
hydrolase activity	1724	121	138.88 -	1.00E+00
transketolase activity	4	0	0.32 -	1.00E+00
transaldolase activity	1	0	0.08 -	1.00E+00
translation regulator activity	126	8	10.15 -	1.00E+00
peroxidase activity	25	0	2.01 -	1.00E+00
structural molecule activity	873	90	70.33 +	1.00E+00
DNA-directed RNA polymerase activity	44	3	3.54 -	1.00E+00
structural constituent of ribosome	194	9	15.63 -	1.00E+00
DNA primase activity	2	0	0.16 -	1.00E+00
tumor necrosis factor receptor activity	1	0	0.08 -	1.00E+00
single-stranded DNA binding	61	7	4.91 +	1.00E+00
nucleotidyltransferase activity	63	5	5.08 -	1.00E+00
double-stranded DNA binding	15	1	1.21 -	1.00E+00
poly(A) RNA binding	58	8	4.67 +	1.00E+00
GTPase activity	115	8	9.26 -	1.00E+00

Figure S2.5 (Continued)

mRNA binding	207	21	16.68 +	1.00E+00
hydrolase activity, hydrolyzing O-glycosyl compounds	32	2	2.58 -	1.00E+00
DNA-directed DNA polymerase activity	28	1	2.26 -	1.00E+00
transforming growth factor beta receptor activity	5	1	0.4 +	1.00E+00
RNA binding	345	31	27.79 +	1.00E+00
neuropeptide hormone activity	1	0	0.08 -	1.00E+00
DNA replication origin binding	35	5	2.82 +	1.00E+00
DNA-methyltransferase activity	19	2	1.53 +	1.00E+00
damaged DNA binding	38	1	3.06 -	1.00E+00
chromatin binding	130	17	10.47 +	1.00E+00
carbohydrate transmembrane transporter activity	87	6	7.01 -	1.00E+00
acetyltransferase activity	81	8	6.53 +	1.00E+00
ion channel activity	256	27	20.62 +	1.00E+00
translation factor activity, nucleic acid binding	149	10	12 -	1.00E+00
transporter activity	730	61	58.81 +	1.00E+00
DNA topoisomerase activity	3	2	0.24 +	1.00E+00
antioxidant activity	29	0	2.34 -	1.00E+00
DNA photolyase activity	2	0	0.16 -	1.00E+00
hormone activity	14	0	1.13 -	1.00E+00
transcription cofactor activity	129	14	10.39 +	1.00E+00
structural constituent of myelin sheath	2	0	0.16 -	1.00E+00
DNA helicase activity	49	1	3.95 -	1.00E+00
amylase activity	16	0	1.29 -	1.00E+00
peptidase inhibitor activity	93	6	7.49 -	1.00E+00
transferase activity, transferring glycosyl groups	176	16	14.18 +	1.00E+00
cyclic nucleotide-gated ion channel activity	9	0	0.73 -	1.00E+00

Figure S2.5 (Continued)

PANTHER Protein Class	Drosophila REFLIST	Target Genes	Target Genes (expected)	over/under	FDR Adj P-value
transcription factor	956	140	77.01	+	1.95E-09
actin family cytoskeletal protein	220	47	17.72	+	7.17E-07
cell adhesion molecule	222	46	17.88	+	2.57E-06
receptor	773	109	62.27	+	3.37E-06
Unclassified	5944	391	478.83	-	8.81E-06
zinc finger transcription factor	312	53	25.13	+	1.08E-04
G-protein modulator	203	39	16.35	+	1.98E-04
cytoskeletal protein	485	71	39.07	+	3.18E-04
voltage-gated sodium channel	12	7	0.97	+	1.20E-02
sodium channel	12	7	0.97	+	1.20E-02
kinase	355	51	28.6	+	1.42E-02
oxidoreductase	712	32	57.36	-	2.51E-02
non-motor actin binding protein	96	20	7.73	+	2.77E-02
guanyl-nucleotide exchange factor	67	16	5.4	+	2.79E-02
calcium channel	15	7	1.21	+	4.65E-02
voltage-gated calcium channel	15	7	1.21	+	4.65E-02
homeobox transcription factor	108	21	8.7	+	4.80E-02
helix-turn-helix transcription factor	108	21	8.7	+	4.80E-02
dehydrogenase	261	7	21.03	-	6.37E-02
protein kinase	230	35	18.53	+	6.57E-02
actin binding motor protein	21	8	1.69	+	6.66E-02
immunoglobulin superfamily cell adhesion molecule	59	14	4.75	+	7.46E-02
nuclear hormone receptor	22	8	1.77	+	9.01E-02
membrane-bound signaling molecule	100	19	8.06	+	1.20E-01
cytokine receptor	36	10	2.9	+	1.51E-01
cell junction protein	50	12	4.03	+	1.70E-01
enzyme modulator	729	83	58.73	+	2.19E-01
membrane trafficking regulatory protein	59	13	4.75	+	2.28E-01
extracellular matrix linker protein	33	9	2.66	+	3.06E-01
immunoglobulin receptor superfamily	29	8	2.34	+	5.05E-01
helicase	99	1	7.98	-	5.43E-01
RNA helicase	66	0	5.32	-	8.72E-01
mitochondrial carrier protein	66	3	5.32	-	1.00E+00
microtubule family cytoskeletal protein	175	16	14.1	+	1.00E+00
microtubule binding motor protein	47	5	3.79	+	1.00E+00
methyltransferase	104	3	8.38	-	1.00E+00
metalloprotease	181	14	14.58	-	1.00E+00
membrane traffic protein	193	26	15.55	+	1.00E+00
major histocompatibility complex antigen	5	1	0.4	+	1.00E+00
mRNA splicing factor	176	18	14.18	+	1.00E+00
mRNA processing factor	207	21	16.68	+	1.00E+00
mRNA polyadenylation factor	58	8	4.67	+	1.00E+00
lyase	163	17	13.13	+	1.00E+00
lipase	112	4	9.02	-	1.00E+00
ligase	367	22	29.56	-	1.00E+00
ligand-gated ion channel	94	7	7.57	-	1.00E+00
kinase modulator	156	10	12.57	-	1.00E+00
kinase inhibitor	63	1	5.08	-	1.00E+00
kinase activator	45	4	3.63	+	1.00E+00
isomerase	158	5	12.73	-	1.00E+00
ionotropic glutamate receptor	51	3	4.11	-	1.00E+00
ion channel	234	25	18.85	+	1.00E+00
intracellular calcium-sensing protein	155	15	12.49	+	1.00E+00
exoribonuclease	25	2	2.01	-	1.00E+00
intermediate filament binding protein	1	0	0.08	-	1.00E+00
exodeoxyribonuclease	12	0	0.97	-	1.00E+00
esterase	108	2	8.7	-	1.00E+00
epimerase/racemase	75	1	6.04	-	1.00E+00
endoribonuclease	45	1	3.63	-	1.00E+00
endodeoxyribonuclease	34	3	2.74	+	1.00E+00
dehydratase	32	2	2.58	-	1.00E+00

Figure S2.5 (Continued)

defense/immunity protein	172	22	13.86 +	1.00E+00
intermediate filament	3	0	0.24 -	1.00E+00
hydroxylase	40	3	3.22 -	1.00E+00
hydrolase	1489	102	119.95 -	1.00E+00
decarboxylase	32	3	2.58 +	1.00E+00
hydratase	19	0	1.53 -	1.00E+00
deaminase	26	3	2.09 +	1.00E+00
deacetylase	20	2	1.61 +	1.00E+00
damaged DNA-binding protein	38	1	3.06 -	1.00E+00
cytokine	12	2	0.97 +	1.00E+00
cysteine protease inhibitor	5	2	0.4 +	1.00E+00
cysteine protease	59	4	4.75 -	1.00E+00
cyclic nucleotide-gated ion channel	9	0	0.73 -	1.00E+00
histone	13	1	1.05 -	1.00E+00
heterotrimeric G-protein	15	1	1.21 -	1.00E+00
guanylate cyclase	26	5	2.09 +	1.00E+00
growth factor	44	6	3.54 +	1.00E+00
glycosyltransferase	167	16	13.45 +	1.00E+00
cyclase	43	8	3.46 +	1.00E+00
glycosidase	49	3	3.95 -	1.00E+00
complement component	13	3	1.05 +	1.00E+00
chromatin/chromatin-binding protein	130	17	10.47 +	1.00E+00
chaperonin	31	1	2.5 -	1.00E+00
chaperone	156	5	12.57 -	1.00E+00
centromere DNA-binding protein	14	2	1.13 +	1.00E+00
glucosidase	11	2	0.89 +	1.00E+00
gap junction	8	1	0.64 +	1.00E+00
galactosidase	5	0	0.4 -	1.00E+00
extracellular matrix structural protein	29	4	2.34 +	1.00E+00
extracellular matrix protein	189	24	15.23 +	1.00E+00
extracellular matrix glycoprotein	48	4	3.87 +	1.00E+00
cation transporter	229	18	18.45 -	1.00E+00
carbohydrate transporter	87	6	7.01 -	1.00E+00
carbohydrate phosphatase	6	0	0.48 -	1.00E+00
carbohydrate kinase	41	2	3.3 -	1.00E+00
calmodulin	155	15	12.49 +	1.00E+00
calcium-binding protein	237	27	19.09 +	1.00E+00
cadherin	13	3	1.05 +	1.00E+00
basic leucine zipper transcription factor	6	3	0.48 +	1.00E+00
basic helix-loop-helix transcription factor	61	9	4.91 +	1.00E+00
aspartic protease	16	1	1.29 -	1.00E+00
apolipoprotein	30	5	2.42 +	1.00E+00
antibacterial response protein	54	3	4.35 -	1.00E+00
annexin	60	3	4.83 -	1.00E+00
anion channel	37	4	2.98 +	1.00E+00
voltage-gated potassium channel	37	1	2.98 -	1.00E+00
amylase	16	0	1.29 -	1.00E+00
voltage-gated ion channel	60	11	4.83 +	1.00E+00
aminoacyl-tRNA synthetase	36	0	2.9 -	1.00E+00
amino acid transporter	106	6	8.54 -	1.00E+00
amino acid kinase	6	0	0.48 -	1.00E+00
aldolase	4	1	0.32 +	1.00E+00
adenylate cyclase	41	8	3.3 +	1.00E+00
acyltransferase	99	7	7.98 -	1.00E+00
viral protein	1	1	0.08 +	1.00E+00
vesicle coat protein	26	5	2.09 +	1.00E+00
ubiquitin-protein ligase	139	9	11.2 -	1.00E+00
tyrosine protein kinase receptor	35	7	2.82 +	1.00E+00
actin and actin related protein	16	3	1.29 +	1.00E+00
acetyltransferase	81	8	6.53 +	1.00E+00
acetylcholine receptor	24	2	1.93 +	1.00E+00
replication origin binding protein	23	3	1.85 +	1.00E+00

Figure S2.5 (Continued)

tumor necrosis factor receptor	1	0	0.08 -	1.00E+00
reductase	231	14	18.61 -	1.00E+00
TGF-beta receptor	5	1	0.4 +	1.00E+00
SNARE protein	22	1	1.77 -	1.00E+00
pyrophosphatase	11	0	0.89 -	1.00E+00
RNA methyltransferase	20	0	1.61 -	1.00E+00
protein phosphatase	104	15	8.38 +	1.00E+00
protein kinase receptor	49	8	3.95 +	1.00E+00
RNA binding protein	808	51	65.09 -	1.00E+00
protease inhibitor	93	6	7.49 -	1.00E+00
protease	627	34	50.51 -	1.00E+00
tumor necrosis factor family member	1	1	0.08 +	1.00E+00
tubulin	12	0	0.97 -	1.00E+00
transporter	857	66	69.04 -	1.00E+00
transmembrane receptor regulatory/adaptor protein	39	4	3.14 +	1.00E+00
translation release factor	6	0	0.48 -	1.00E+00
translation initiation factor	74	8	5.96 +	1.00E+00
translation factor	128	10	10.31 -	1.00E+00
KRAB box transcription factor	114	18	9.18 +	1.00E+00
translation elongation factor	64	2	5.16 -	1.00E+00
Hsp90 family chaperone	3	0	0.24 -	1.00E+00
transketolase	4	0	0.32 -	1.00E+00
Hsp70 family chaperone	12	0	0.97 -	1.00E+00
primase	2	0	0.16 -	1.00E+00
transferase	983	96	79.19 +	1.00E+00
potassium channel	37	1	2.98 -	1.00E+00
phosphorylase	9	0	0.73 -	1.00E+00
phospholipase	22	3	1.77 +	1.00E+00
HMG box transcription factor	31	3	2.5 +	1.00E+00
GABA receptor	24	2	1.93 +	1.00E+00
phosphodiesterase	26	4	2.09 +	1.00E+00
phosphatase modulator	35	2	2.82 -	1.00E+00
G-protein coupled receptor	241	25	19.41 +	1.00E+00
phosphatase inhibitor	14	2	1.13 +	1.00E+00
G-protein	115	8	9.26 -	1.00E+00
phosphatase activator	3	0	0.24 -	1.00E+00
phosphatase	221	21	17.8 +	1.00E+00
peroxidase	25	0	2.01 -	1.00E+00
transfer/carrier protein	294	18	23.68 -	1.00E+00
transcription cofactor	129	14	10.39 +	1.00E+00
transaminase	17	4	1.37 +	1.00E+00
transaldolase	1	0	0.08 -	1.00E+00
tight junction	5	1	0.4 +	1.00E+00
DNA-directed RNA polymerase	42	3	3.38 -	1.00E+00
surfactant	11	1	0.89 +	1.00E+00
structural protein	196	14	15.79 -	1.00E+00
DNA-directed DNA polymerase	28	1	2.26 -	1.00E+00
storage protein	47	1	3.79 -	1.00E+00
peptide hormone	14	0	1.13 -	1.00E+00
DNA topoisomerase	3	2	0.24 +	1.00E+00
DNA strand-pairing protein	5	0	0.4 -	1.00E+00
oxygenase	151	4	12.16 -	1.00E+00
DNA polymerase processivity factor	3	0	0.24 -	1.00E+00
DNA photolyase	2	0	0.16 -	1.00E+00
oxidase	122	5	9.83 -	1.00E+00
DNA methyltransferase	19	2	1.53 +	1.00E+00
nucleotidyltransferase	63	5	5.08 -	1.00E+00
DNA ligase	4	0	0.32 -	1.00E+00
DNA helicase	49	1	3.95 -	1.00E+00
nucleotide phosphatase	39	1	3.14 -	1.00E+00
nucleotide kinase	48	5	3.87 +	1.00E+00
DNA glycosylase	4	1	0.32 +	1.00E+00

Figure S2.5 (Continued)

nucleic acid binding	1761	162	141.86 +	1.00E+00
nuclease	172	15	13.86 +	1.00E+00
small GTPase	68	4	5.48 -	1.00E+00
signaling molecule	442	51	35.61 +	1.00E+00
serine/threonine protein kinase receptor	23	2	1.85 +	1.00E+00
serine protease inhibitor	62	4	4.99 -	1.00E+00
serine protease	349	16	28.11 -	1.00E+00
ribosomal protein	194	9	15.63 -	1.00E+00
DNA binding protein	624	62	50.27 +	1.00E+00
ribonucleoprotein	79	13	6.36 +	1.00E+00
non-receptor tyrosine protein kinase	51	7	4.11 +	1.00E+00
CREB transcription factor	6	3	0.48 +	1.00E+00
non-receptor serine/threonine protein kinase	147	18	11.84 +	1.00E+00
non-motor microtubule binding protein	51	7	4.11 +	1.00E+00
ATP-binding cassette (ABC) transporter	87	4	7.01 -	1.00E+00
ATP synthase	53	4	4.27 -	1.00E+00
neuropeptide	1	0	0.08 -	1.00E+00
myelin protein	2	0	0.16 -	1.00E+00
mutase	8	0	0.64 -	1.00E+00

The PANTHER gene ontology annotation database was used to identify protein class and molecular function annotations overrepresented in the set of genes targeted by differentially expressed microRNAs.(*4*)

Figure S2.6. Diana MicroT target predictions for genes with "Neurological systems process" GO annotation

FlyBase ID	Symbol	Name	dme-miR-312-3p	dme-miR-314-3p	dme-miR-956-3p	dme-miR-958-3p	dme-miR-958-5p
FBgn0004396	CrebA	Cyclic-AMP response element binding protein A		1	0	0	0	0
FBgn0038439	Cad89D	Cadherin 89D		1	0	0	0	0
FBgn0029835	CG5921		0	1	0	0	0	0
FBgn0017549		0	0	1	1	0	0	0
FBgn0039530	Tusp	Tusp		1	0	0	0	0
FBgn0086475	sec3	sec3		1	0	0	0	0
FBgn0034509	Obp57c	Odorant-binding protein 57c		1	0	0	0	0
FBgn0039431	CG6490		0	1	0	1	0	0
FBgn0003870	ttk	tramtrack		1	0	1	0	0
FBgn0262737	mub	mushroom-body expressed		1	0	1	0	0
FBgn0029846	#N/A	#N/A		1	0	0	0	0
FBgn0259750	#N/A	#N/A		1	0	0	0	0
FBgn0039380	CG5890		0	1	0	0	0	0
FBgn0035308	CG15822		0	1	0	0	0	0
FBgn0015323	VAChT	VAChT		1	0	0	0	0
FBgn0039187	CG6454		0	1	0	0	0	0
FBgn0004914	Hnf4	Hepatocyte nuclear factor 4		1	0	0	0	0
FBgn0013759	CASK	CASK ortholog		1	0	0	0	0
FBgn0085414	dpr12	dpr12		1	0	0	0	0
FBgn0015129		0	0	1	0	0	0	0
FBgn0035540	Syx17	Syntaxin 17		1	0	0	0	0
FBgn0052600	dpr8	dpr8		1	0	0	0	0
FBgn0028872	CG18095		0	1	0	0	0	0
FBgn0000120	Arr1	Arrestin 1		1	0	0	0	0
FBgn0262350	#N/A	#N/A		1	1	0	0	0
FBgn0033138	Tsp42Eq	Tetraspanin 42Eq		1	0	0	0	0
FBgn0039528	dsd	distracted		1	0	0	0	0
FBgn0014870	Psi	P-element somatic inhibitor		1	0	0	0	1
FBgn0034433	endoB	endophilin B		1	0	0	1	0
FBgn0087007	bbg	big bang		1	0	0	0	0
FBgn0003149	Prm	Paramyosin		1	0	0	0	0
FBgn0259231	CCKLR-17D1	CCK-like receptor at 17D1		1	1	0	0	0
FBgn0024944	Oamb	Octopamine receptor in mushroom bodies		1	0	0	0	0
FBgn0036927	CG7433		0	1	0	0	0	0
FBgn0004242	Syt1	Synaptotagmin 1		1	1	0	0	0
FBgn0039911	CG1909		0	1	0	0	0	0
FBgn0000037	mAcR	muscarinic Acetylcholine Receptor		1	0	0	0	0
FBgn0261836	Msp-300	Muscle-specific protein 300		1	0	0	0	0
FBgn0040395	Unc-76	Unc-76		1	0	0	0	0
FBgn0000210	br	broad		1	0	0	0	0
FBgn0000317	ck	crinkled		1	0	0	0	0
FBgn0250910	Octbeta3R	Octopamine beta3 receptor		1	0	0	0	0
FBgn0085383	CG34354		0	1	0	0	0	0
FBgn0031981	CG7466		0	1	0	0	0	0
FBgn0040334	Tsp3A	Tetraspanin 3A		1	0	0	0	0
FBgn0038339	CG6118		0	1	0	0	0	0
FBgn0028875	nAcRalpha-34E	nicotinic Acetylcholine Receptor alpha 34E		1	0	0	0	0
FBgn0028371	jbug	jitterbug		1	0	0	0	0
FBgn0015295	shark	SH2 ankyrin repeat kinase		1	0	0	0	0
FBgn0031760	Tsp26A	Tetraspanin 26A		1	0	0	1	0
FBgn0037993	dpr15	dpr15		1	0	0	0	0
FBgn0028704	Nckx30C	Nckx30C		1	1	0	0	0
FBgn0032341	Reps		0	1	0	1	1	0
FBgn0033058	CCHa2r	CCHamide-2 receptor		1	0	0	0	0
FBgn0001297	kay	kayak		1	0	0	0	0
FBgn0026386	Or47a	Odorant receptor 47a		1	0	0	0	0
FBgn0028582	lqf	liquid facets		1	0	0	0	0
FBgn0037326	CG14669		0	1	0	0	0	0
FBgn0040208	Kat60	Katanin 60		1	0	0	0	0
FBgn0052702	CG32702		0	1	0	0	0	0
FBgn0013733	shot	short stop		1	0	0	1	0
FBgn0261552	ps	pasilla		1	0	0	0	0
FBgn0000273	Pka-C1	cAMP-dependent protein kinase 1		1	0	0	0	0
FBgn0085434	NaCP60E	Na channel protein 60E		1	0	0	0	0
FBgn0033876	Syngr	Synaptogyrin		1	0	0	0	0
FBgn0015774	NetB	Netrin-B		1	0	0	1	0
FBgn0262111	f	forked		1	0	0	0	0
FBgn0010473	tutl	turtle		1	0	0	0	0
FBgn0035170	dpr20	dpr20		1	0	0	0	0
FBgn0261999	CG42817		0	1	0	0	0	0
FBgn0037848	Tsp86D	Tetraspanin 86D		1	0	0	0	0
FBgn0029508	Tsp42Ea	Tetraspanin 42Ea		1	0	0	1	0
FBgn0261387	CG17528		0	1	0	0	0	0
FBgn0024277	trio	trio		1	0	0	0	0
FBgn0037130	Syn1	Syntrophin-like 1		1	0	0	0	0
FBgn0028431	#N/A	#N/A		1	0	0	0	0
FBgn0038237	Pde6	Phosphodiesterase 6		1	0	0	0	0

Figure S2.6 (Continued)

FBgn0085405	CG34376		0	1	0	0	0	0
FBgn0038880	SIFR	SIFamide receptor		1	0	0	0	0
FBgn0011582	Dop1R1	Dopamine 1-like receptor 1		1	1	1	1	0
FBgn0000448	Hr46	Hormone receptor-like in 46		1	1	0	0	0
FBgn0023000	mth	methuselah		1	0	0	0	0
FBgn0053516	dpr3	dpr3		1	0	0	0	0
FBgn0026086	Adar	Adenosine deaminase acting on RNA		1	0	1	0	0
FBgn0023091	dimm	dimmed		1	0	0	0	0
FBgn0053555	#N/A	#N/A		0	1	0	0	0
FBgn0040388	boi	brother of ihog		0	1	1	0	0
FBgn0038118	timeout	timeout		0	1	0	0	0
FBgn0263218	0		0	0	1	1	0	0
FBgn0034691	synj	synaptojanin		0	1	0	0	0
FBgn0085397	Fili	Fish-lips		0	1	0	0	0
FBgn0025632	CG4313		0	0	1	0	0	0
FBgn0004622	Takr99D	Tachykinin-like receptor at 99D		0	1	1	0	0
FBgn0037408	NPFR1	neuropeptide F receptor		0	1	0	0	0
FBgn0028550	Atf3	Activating transcription factor 3		0	1	1	0	0
FBgn0016059	Sema-1b	Sema-1b		0	1	1	0	0
FBgn0000667	Actn	alpha actinin		0	1	0	0	0
FBgn0086783	#N/A	#N/A		0	1	0	0	0
FBgn0003206	Ras64B	Ras oncogene at 64B		0	1	0	0	0
FBgn0034135	Syn2	Syntrophin-like 2		0	1	0	0	0
FBgn0024963	GluClalpha	GluClalpha		0	1	0	0	1
FBgn0030766	mthl1	methuselah-like 1		0	1	0	0	0
FBgn0002940	ninaE	neither inactivation nor afterpotential E		0	1	0	0	0
FBgn0261053	Cad86C	Cadherin 86C		0	1	0	0	0
FBgn0039335	Vps33B	Vacuolar protein sorting 33B		0	1	0	0	0
FBgn0034049	bdg	bedraggled		0	1	0	0	0
FBgn0038720	CG6231		0	0	1	0	0	0
FBgn0038139	PK2-R2	Pyrokinin 2 receptor 2		0	1	0	0	0
FBgn0000633	fas	faint sausage		0	1	0	0	0
FBgn0261085	Syt12	Synaptotagmin 12		0	1	0	0	0
FBgn0038089	d-cup	davis-cup		0	0	1	0	0
FBgn0036260	Rh7	Rhodopsin 7		0	0	1	0	0
FBgn0029974	dpr14	dpr14		0	0	1	0	0
FBgn0038890	CG7956		0	0	0	1	1	0
FBgn0039157	Myo95E	Myosin 95E		0	0	1	0	0
FBgn0000635	Fas2	Fasciclin 2		0	0	1	0	0
FBgn0053202	dpr11	dpr11		0	0	1	0	0
FBgn0031119	CG1812		0	0	0	1	0	0
FBgn0259225	#N/A	#N/A		0	0	1	0	0
FBgn0263102	psq	pipsqueak		0	0	1	0	0
FBgn0085446	CG34417		0	0	0	1	1	0
FBgn0034286	dpr13	dpr13		0	0	1	0	0
FBgn0011225	jar	jaguar		0	0	1	0	0
FBgn0262742	Fas1	Fasciclin 1		0	0	1	0	0
FBgn0039069	CG6763		0	0	0	1	0	0
FBgn0039927	CG11155		0	0	0	1	0	0
FBgn0003861	trp	transient receptor potential		0	0	0	1	0
FBgn0002931	net	net		0	0	0	1	0
FBgn0086365	Orct2	Organic cation transporter 2		0	0	0	1	0
FBgn0036273	CG10426		0	0	0	0	1	0
FBgn0033679	CG8888		0	0	0	0	1	0
FBgn0039644	CG11897		0	0	0	0	1	0
FBgn0037552	CG7800		0	0	0	0	1	0
FBgn0036789	AlCR2	allatostatin C receptor 2		0	0	0	1	0
FBgn0000346	comt	comatose		0	0	0	1	0
FBgn0038826	Syp	Syncrip		0	0	0	1	0
FBgn0259822	Ca-beta	Ca2+-channel-protein-beta-subunit		0	0	0	1	0
FBgn0011259	Sema-1a	Sema-1a		0	0	0	1	0
FBgn0015609	CadN	Cadherin-N		0	0	0	1	0
FBgn0028433	Ggamma30A	G protein gamma30A		0	0	0	1	0
FBgn0010399	Nmdar1	NMDA receptor 1		0	0	0	1	0
FBgn0261090	Sytbeta	Synaptotagmin beta		0	0	0	1	0
FBgn0033739	Dyb	Dystrobrevin-like		0	0	0	1	0
FBgn0028400	Syt4	Synaptotagmin 4		0	0	0	0	1
FBgn0029762	NAAT1	Nutrient Amino Acid Transporter 1		0	0	0	0	1
FBgn0085429	#N/A	#N/A		0	0	0	0	1
FBgn0041605	cpx	complexin		0	0	0	0	1
FBgn0025631	moody	moody		0	0	0	0	1
FBgn0036381	CG8745		0	0	0	0	0	1
FBgn0020258	ppk	pickpocket		0	0	0	0	1
FBgn0011666	msi	musashi		0	0	0	0	1
FBgn0003651	svp	seven up		0	0	0	0	1
FBgn0259246	brp	bruchpilot		0	0	0	0	1

The PANTHER gene ontology annotation database was used to identify biological process gene ontology annotations overrepresented in the set of genes targeted by differentially expressed microRNAs.(*4*) This analysis identifies the "Neurological systems process" as an overrepresented annotation. Due to our interest in identifying target genes of differentially expressed microRNAs that are involved in memory formation, this term was used to refine the results of our PANTHER analysis.

Figure S3.1. High exon coverage transcripts

HECT	Symbol	plus_reads	plus_coverage	minus_reads	minus_coverage	plus_minus_read_ratio	intron
FBtr0070040	CG32230	1626	0.9817768	4	0.1321184	406.5	2
FBtr0070041	CG32230	1884	1	3	0.1134565	628	3
FBtr0070156	RpL36	3539	0.95	31	0.2982759	114.1612903	2
FBtr0070159	CG13364	272	0.9895833	11	0.203125	24.72727273	1
FBtr0070179	CG14629	1363	0.9774368	29	0.2057762	47	0
FBtr0070381	CG3835	2175	0.9769065	59	0.1804511	36.86440678	4
FBtr0070386	CG3621	1338	0.9764706	26	0.1058824	51.46153846	1
FBtr0070611	VhaAC39-1	1636	0.9526726	47	0.185412	34.80851064	2
FBtr0070708	l(1)G0334	7994	0.9817981	39	0.1395477	204.974359	4
FBtr0070709	l(1)G0334	7595	0.9820036	36	0.1421716	210.9722222	5
FBtr0070801	RpL35	3145	0.9983444	39	0.2996689	80.64102564	2
FBtr0070822	Act5C	7981	0.9743433	32	0.1954795	249.40625	1
FBtr0070907	kdn	24324	0.9911802	53	0.2225955	458.9433962	2
FBtr0070909	Marf	3642	0.9686684	81	0.2106179	44.96296296	7
FBtr0070910	Marf	3836	0.9634146	84	0.2103658	45.66666667	7
FBtr0070915	RpL7A	6487	0.9802073	51	0.2196041	127.1960784	5
FBtr0070916	RpL7A	6530	0.9725023	41	0.2227314	159.2682927	4
FBtr0070924	CG3446	1443	0.9654812	37	0.3483264	39	2
FBtr0070933	Ubi-p5E	13130	0.9796239	99	0.2032393	132.6262626	1
FBtr0070953	l(1)G0255	6464	0.9894552	24	0.0849443	269.3333333	4
FBtr0070954	l(1)G0255	6137	0.9611536	32	0.0971159	191.78125	4
FBtr0071094	RpS14a	3075	0.9875583	31	0.3499222	99.19354839	2
FBtr0071123	CG2233	7862	0.9810811	27	0.1405405	291.1851852	4
FBtr0071135	RpS6	3815	0.9927235	19	0.1528067	200.7894737	2
FBtr0071140	CG18624	1919	0.9876289	6	0.1237113	319.8333333	1
FBtr0071180	ND75	9497	0.9938295	53	0.1542615	179.1886792	5
FBtr0071181	ND75	9341	0.98663	54	0.1631931	172.9814815	4
FBtr0071345	His3.3B	1720	0.9562764	96	0.2165021	17.91666667	2
FBtr0071360	RpS28b	3529	0.975048	11	0.2072937	320.8181818	1
FBtr0071361	Hex-A	4594	0.9534236	66	0.1926752	69.60606061	0
FBtr0071362	Hex-A	4594	0.9534236	66	0.1926752	69.60606061	0
FBtr0071419	Yp1	38518	0.9961783	67	0.3025478	574.8955224	1
FBtr0071424	Yp2	29336	0.9967804	173	0.3251771	169.5722543	1
FBtr0071438	l(1)G0230	2825	0.9924925	17	0.2117117	166.1764706	3
FBtr0071444	CG17841	1594	0.9546729	37	0.1565421	43.08108108	3
FBtr0071449	Neb-cGP	2638	0.9926829	15	0.204878	175.8666667	2
FBtr0071498	Atg8a	3126	0.9636664	85	0.4459125	36.77647059	2
FBtr0071519	Act57B	10459	0.9831982	97	0.1755504	107.8247423	1
FBtr0071537	Treh	3529	0.961929	135	0.2407542	26.14074074	5
FBtr0071540	Treh	3503	0.9596031	151	0.2625797	23.1986755	5
FBtr0071592	RpL29	1164	0.9908537	10	0.222561	116.4	3
FBtr0071593	RpL29	1183	1	10	0.1931217	118.3	2
FBtr0071599	CG9485	5031	0.9508163	257	0.2367347	19.57587549	7
FBtr0071601	CG9485	5038	0.9555146	261	0.2443624	19.30268199	7
FBtr0071663	CG10320	1925	0.9533169	5	0.1228501	385	1
FBtr0071784	Swim	2512	0.951148	45	0.2188569	55.82222222	3
FBtr0071785	Swim	2783	0.9647533	42	0.2019134	66.26190476	3
FBtr0071813	Gp150	5968	0.973983	110	0.1903974	54.25454545	5
FBtr0071814	Gp150	5972	0.9668703	111	0.1926692	53.8018018	5
FBtr0071815	Gp150	5958	0.9616533	118	0.1986223	50.49152542	4
FBtr0071855	RpS16	4156	0.9740933	14	0.1830743	296.8571429	4
FBtr0071883	blw	40834	0.9982571	93	0.1350763	439.0752688	3
FBtr0071897	RpL23	2136	0.9797794	9	0.1727941	237.3333333	3
FBtr0071935	RpS24	2130	0.9860465	8	0.1534884	266.25	1
FBtr0072030	CG30415	3047	0.9741518	32	0.2875606	95.21875	2
FBtr0072031	CG30415	2769	0.9695586	35	0.3105023	79.11428571	1
FBtr0072061	levy	4834	0.9712838	13	0.1334459	371.8461538	1
FBtr0072121	CG3906	1351	0.9941107	16	0.147232	84.4375	1
FBtr0072141	Tal	1221	0.9604672	27	0.1132075	45.22222222	2
FBtr0072164	PebIII	2149	0.9698997	66	0.4481605	32.56060606	1
FBtr0072172	eIF-5A	2933	0.9870634	76	0.1746442	38.59210526	4
FBtr0072173	eIF-5A	3338	0.98713	76	0.1737452	43.92105263	3
FBtr0072175	RpL12	3428	0.9633758	19	0.2085987	180.4210526	4
FBtr0072176	RpL12	3283	0.9595016	19	0.2040498	172.7894737	3
FBtr0072185	RpL39	3991	0.9895105	8	0.3286713	498.875	2
FBtr0072188	tsr	924	0.9571984	15	0.1258106	61.6	3

Figure S3.1 (Continued)

FBtr0072211	Ca-P60A	29667	0.99694	111	0.1704406	267.2702703	8
FBtr0072212	Ca-P60A	30267	0.9921781	114	0.1729844	265.5	8
FBtr0072213	Ca-P60A	29377	0.9872856	111	0.1646954	264.6576577	8
FBtr0072214	Ca-P60A	29338	0.9905344	120	0.1822901	244.4833333	8
FBtr0072216	Ca-P60A	32431	0.9610364	187	0.1729982	173.4278075	8
FBtr0072217	Ca-P60A	29509	0.9814073	122	0.1893593	241.8770492	8
FBtr0072339	CG4692	2522	0.9744526	16	0.2372263	157.625	2
FBtr0072340	CG4692	2087	0.9965096	17	0.2547993	122.7647059	1
FBtr0072405	RpL19	3874	0.9697352	25	0.2055486	154.96	2
FBtr0072406	RpL19	3873	0.9960053	21	0.2050599	184.4285714	2
FBtr0072436	GstE12	2121	0.9875	67	0.3840909	31.65671642	2
FBtr0072676	mtacp1	2236	0.9716216	30	0.1594595	74.53333333	3
FBtr0072677	mtacp1	1924	0.9824	18	0.1408	106.8888889	3
FBtr0072805	RpL23A	4643	0.9872047	34	0.1486221	136.5588235	2
FBtr0072848	sls	33663	0.9726844	655	0.1607494	51.39389313	13
FBtr0072924	RpL8	3834	0.993988	18	0.1342685	213	2
FBtr0073040	Hsp83	3694	0.9767784	112	0.237007	32.98214286	1
FBtr0073097	RpL28	2101	0.9798903	14	0.2522852	150.0714286	3
FBtr0073113	CG12079	3047	0.9979508	26	0.2858607	117.1923077	2
FBtr0073151	Scsalpha	3048	0.9858012	44	0.2082488	69.27272727	3
FBtr0073295	DOR	5948	0.9727391	146	0.1389009	40.73972603	4
FBtr0073296	DOR	5018	0.9737883	145	0.1716123	34.60689655	4
FBtr0073421	sesB	40973	0.998164	84	0.2331701	487.7738095	3
FBtr0073423	sesB	42688	0.9969512	86	0.2567073	496.372093	3
FBtr0073439	CG15201	517	0.9672801	14	0.200409	36.92857143	2
FBtr0073452	CG11752	620	0.9697581	48	0.578629	12.91666667	0
FBtr0073495	Gs2	5689	0.9758375	243	0.4387699	23.41152263	4
FBtr0073496	Gs2	5432	0.9756772	243	0.4416805	22.35390947	4
FBtr0073539	CG1561	3321	0.9884259	86	0.2106481	38.61627907	1
FBtr0073576	regucalcin	1753	0.9537205	42	0.2994555	41.73809524	2
FBtr0073763	Jafrac1	1042	0.9657258	34	0.3487903	30.64705882	1
FBtr0073792	RpS15Aa	3996	0.9548105	19	0.3119534	210.3157895	2
FBtr0073793	RpS15Aa	3956	0.983631	18	0.2946429	219.7777778	2
FBtr0073794	RpS15Aa	3987	0.9510204	24	0.3102041	166.125	1
FBtr0073795	RpS15Aa	3955	0.9876733	18	0.3050847	219.7222222	1
FBtr0073821	Yp3	30772	1	43	0.2619208	715.627907	2
FBtr0073851	up	12349	0.9881735	17	0.130749	726.4117647	9
FBtr0073852	up	12331	0.987984	16	0.1228304	770.6875	8
FBtr0073960	CG5548	832	0.9897959	6	0.1292517	138.6666667	1
FBtr0073979	CG9512	3289	0.9733967	56	0.1857482	58.73214286	2
FBtr0074112	Gapdh2	8218	0.9646569	76	0.4026334	108.1315789	1
FBtr0074113	#N/A	8214	0.9855272	86	0.4162647	95.51162791	#N/A
FBtr0074151	CG9172	2462	0.9955899	19	0.1102536	129.5789474	1
FBtr0074152	CG9172	2435	0.9754738	15	0.1125975	162.3333333	0
FBtr0074193	CG8952	775	0.9645233	17	0.1618625	45.58823529	1
FBtr0074311	RpS19a	3023	0.9637795	32	0.1874016	94.46875	2
FBtr0074312	RpS19a	3032	0.9649924	32	0.1811263	94.75	2
FBtr0074387	CG5010	1985	0.9510416	25	0.2270833	79.4	1
FBtr0074406	RpS5a	3785	0.9964913	14	0.1532164	270.3571429	4
FBtr0074520	wupA	5112	0.987931	298	0.1689655	17.15436242	7
FBtr0074521	wupA	5091	0.987931	299	0.1818966	17.02675585	7
FBtr0074522	wupA	5621	0.987931	298	0.1689655	18.86241611	7
FBtr0074523	wupA	5067	0.9672414	298	0.1689655	17.0033557	7
FBtr0074524	wupA	5990	0.9895756	300	0.1571109	19.96666667	8
FBtr0074559	Tsf1	4201	0.9890329	191	0.3933201	21.9947644	4
FBtr0074626	CG15043	3586	0.9971057	15	0.1837916	239.0666667	1
FBtr0074712	CG12203	1572	0.9909209	23	0.151751	68.34782609	2
FBtr0074731	RpS10b	2616	0.974428	17	0.243607	153.8823529	2
FBtr0074732	RpS10b	2571	0.9903846	16	0.2266483	160.6875	2
FBtr0074780	l(1)G0156	11205	0.9887387	114	0.545045	98.28947368	4
FBtr0074815	CoVIb	4839	0.9850948	32	0.3699187	151.21875	1
FBtr0074816	CoVIb	4911	0.9970194	30	0.3457526	163.7	1
FBtr0074817	CoVIb	4912	0.9971265	36	0.3548851	136.4444444	1
FBtr0074910	fln	7306	0.9908814	28	0.331307	260.9285714	3
FBtr0075014	Gbs-76A	5241	0.9517086	75	0.1440271	69.88	2
FBtr0075043	CG18135	3877	0.9580958	44	0.0885015	88.11363636	1
FBtr0075048	CG3819	2082	0.9593679	27	0.1392024	77.11111111	1

Figure S3.1 (Continued)

FBtr0075058	Cat	3558	0.9613648	98	0.2112393	36.30612245	2
FBtr0075066	RpL26	3148	0.9922839	22	0.2283951	143.0909091	1
FBtr0075069	CG6839	13708	0.9879245	24	0.0935849	571.1666667	1
FBtr0075157	CG5506	1027	0.9591528	30	0.1694402	34.23333333	1
FBtr0075217	CG7630	2175	0.9976019	19	0.1870504	114.4736842	3
FBtr0075220	CG7603	1744	0.9645669	19	0.2191601	91.78947368	1
FBtr0075263	Nc73EF	17298	0.96095	131	0.1461521	132.0458015	10
FBtr0075264	Nc73EF	16723	0.9829637	117	0.1446728	142.9316239	10
FBtr0075266	Nc73EF	17469	0.9668798	142	0.1617177	123.0211268	10
FBtr0075267	Nc73EF	16574	0.9938055	116	0.1400485	142.8793103	10
FBtr0075268	Nc73EF	16841	0.9937602	117	0.1451438	143.9401709	10
FBtr0075269	Nc73EF	16599	0.9933137	122	0.1473656	136.057377	10
FBtr0075341	CG9674	5378	0.9612254	164	0.1087761	32.79268293	14
FBtr0075343	CG9674	2532	0.9736414	24	0.109014	105.5	5
FBtr0075369	Nplp3	1648	1	16	0.2379182	103	1
FBtr0075375	retinin	2068	0.9862155	6	0.1165414	344.6666667	1
FBtr0075426	Pdh	2852	0.9989418	16	0.1216931	178.25	2
FBtr0075427	Pdh	2822	0.9989384	16	0.1220807	176.375	2
FBtr0075492	Pgm	3408	0.9945741	167	0.4910472	20.40718563	3
FBtr0075648	Pdi	2882	0.9807692	42	0.1591346	68.61904762	1
FBtr0075718	Mpcp	4939	0.9894118	103	0.3088235	47.95145631	3
FBtr0075719	Mpcp	4759	0.9641435	103	0.2988048	46.2038835	3
FBtr0075839	Nplp2	3977	0.9933628	7	0.1172566	568.1428571	2
FBtr0075878	RpS12	2682	0.9722675	23	0.1843393	116.6086957	3
FBtr0075884	RpS4	5093	0.9508197	22	0.1639344	231.5	5
FBtr0076003	Est-6	2126	0.9655374	18	0.1063084	118.1111111	1
FBtr0076032	RpL10Ab	5586	0.984	23	0.2106667	242.8695652	2
FBtr0076044	CG14125	7036	0.9976162	35	0.1418355	201.0285714	1
FBtr0076119	Muc68D	57282	0.9920866	234	0.2132445	244.7948718	1
FBtr0076229	Sod	785	0.9687075	28	0.1306122	28.03571429	1
FBtr0076389	CG18180	8019	0.9988901	38	0.3540511	211.0263158	0
FBtr0076405	ATPsyn-b	9323	0.9934641	25	0.167756	372.92	2
FBtr0076423	RpS9	6078	0.9904632	14	0.2411444	434.1428571	3
FBtr0076425	RpS9	5932	0.9957507	14	0.2507082	423.7142857	2
FBtr0076479	RpS17	2463	0.9738318	20	0.3046729	123.15	3
FBtr0076501	UGP	4635	0.9557564	147	0.2630828	31.53061224	8
FBtr0076530	CG13315	3412	0.9935588	19	0.3494364	179.5789474	0
FBtr0076545	Argk	14341	0.9993417	34	0.2139565	421.7941176	1
FBtr0076593	Prm	10554	0.9795138	96	0.1761814	109.9375	8
FBtr0076594	Prm	9653	0.983881	95	0.1949783	101.6105263	8
FBtr0076595	Prm	8759	0.9641838	92	0.1464342	95.20652174	4
FBtr0076596	Prm	7858	0.9669118	91	0.1639706	86.35164835	4
FBtr0076599	Arr2	4567	0.989011	63	0.1215255	72.49206349	2
FBtr0076633	RpL14	3617	0.9838449	12	0.1260097	301.4166667	3
FBtr0076667	Idh	4251	0.9845297	21	0.1280941	202.4285714	1
FBtr0076668	Idh	4249	0.9708384	22	0.1348724	193.1363636	1
FBtr0076808	CG12262	1587	0.9589322	141	0.5509925	11.25531915	2
FBtr0076892	RpL18	2458	0.9903315	49	0.3812155	50.16326531	3
FBtr0077038	yip7	8467	0.9955752	53	0.449115	159.754717	1
FBtr0077040	Jon65Aiv	15113	0.9856195	19	0.1570797	795.4210526	1
FBtr0077041	Jon65Aiii	9824	0.9977143	6	0.0857143	1637.333333	1
FBtr0077131	Msr-110	5422	0.961058	90	0.213446	60.24444444	6
FBtr0077132	Msr-110	5413	0.9606679	90	0.2155844	60.14444444	5
FBtr0077144	CG4769	10269	0.9953977	43	0.2629849	238.8139535	5
FBtr0077304	Obp19d	3990	0.9968102	15	0.2982456	266	3
FBtr0077431	Tps1	4855	0.9859589	150	0.2828767	32.36666667	4
FBtr0077452	RpL27A	3508	0.9503425	6	0.0839041	584.6666667	3
FBtr0077470	RpL40	3272	0.9693356	50	0.4514481	65.44	1
FBtr0077520	Pdsw	2439	0.992163	11	0.2507837	221.7272727	3
FBtr0077521	Pdsw	2313	0.9658915	13	0.2372093	177.9230769	2
FBtr0077524	Thor	1785	0.9735099	153	0.3536424	11.66666667	1
FBtr0077550	CG16704	176	0.9791667	15	0.1428571	11.73333333	1
FBtr0077617	CG12400	1770	0.9781659	21	0.2008734	84.28571429	2
FBtr0077621	RpS21	2207	0.9693094	1	0.0383632	2207	3
FBtr0077659	CG3523	16629	0.9845743	336	0.1816284	49.49107143	5
FBtr0077739	Pgk	6363	0.9979811	60	0.2469717	106.05	2
FBtr0077828	GlyP	14988	0.9977735	126	0.2490458	118.952381	3

Figure S3.1 (Continued)

FBtr0077867	Got2	2065	0.9542857	53	0.2101587	38.96226415	4
FBtr0077909	Eno	18551	0.995439	80	0.2035348	231.8875	1
FBtr0077915	RFeSP	3922	0.9882491	20	0.226792	196.1	2
FBtr0078056	RpLP1	3120	0.9950083	27	0.3760399	115.5555556	1
FBtr0078118	CG11455	934	0.976	86	0.592	10.86046512	1
FBtr0078154	CG3164	2968	0.9521238	55	0.1691865	53.96363636	7
FBtr0078382	CoVIII	1608	0.997537	3	0.0369458	536	1
FBtr0078477	CG7470	3669	0.9745467	74	0.1666667	49.58108108	11
FBtr0078481	RpLP0	5509	0.9925681	77	0.1907514	71.54545455	2
FBtr0078643	CG1213	1460	0.9591141	51	0.1607042	28.62745098	3
FBtr0078650	Rm62	2195	0.95321	170	0.20856	12.91176471	5
FBtr0078655	Obp83b	204	0.954717	18	0.154717	11.33333333	2
FBtr0078662	CG2017	1791	0.9553528	63	0.2568411	28.42857143	3
FBtr0078700	Vha26	3261	0.972363	63	0.2851221	51.76190476	4
FBtr0078701	Vha26	3263	0.9616571	64	0.2794183	50.984375	3
FBtr0078705	RpL13A	4157	0.9750692	18	0.131579	230.9444444	2
FBtr0078745	exba	2446	0.9708798	32	0.185254	76.4375	7
FBtr0078769	RpL35A	2161	0.9565217	12	0.1723027	180.0833333	4
FBtr0078822	CG12163	1695	0.9585547	37	0.1859724	45.81081081	4
FBtr0078908	CG14645	3367	1	16	0.2518337	210.4375	0
FBtr0078968	Gel	2959	0.954177	50	0.1240747	59.18	7
FBtr0078973	Gel	2908	0.9540598	50	0.1253561	58.16	7
FBtr0079001	Cg25C	3926	0.9625555	145	0.1167702	27.07586207	8
FBtr0079003	Cg25C	4050	0.9541667	153	0.1205	26.47058824	8
FBtr0079016	RpL37A	2420	0.9706458	22	0.2896282	110	3
FBtr0079025	CG8680	612	0.9779412	13	0.2224265	47.07692308	2
FBtr0079031	eIF-3p40	1135	0.9729272	60	0.4373943	18.91666667	2
FBtr0079053	Trip1	843	0.9660153	37	0.2166525	22.78378378	2
FBtr0079055	Jon25Bii	4541	0.9920815	10	0.1131222	454.1	0
FBtr0079056	Jon25Biii	3748	0.9975816	27	0.2756953	138.8148148	0
FBtr0079072	cype	2504	0.9947644	7	0.2486911	357.7142857	2
FBtr0079147	Gpdh	13269	0.9963213	40	0.202943	331.725	6
FBtr0079175	eIF-4a	4855	0.9668835	45	0.1738616	107.8888889	4
FBtr0079176	eIF-4a	4835	0.9819704	44	0.1790084	109.8863636	5
FBtr0079177	eIF-4a	5106	0.9593936	52	0.1781267	98.19230769	4
FBtr0079178	eIF-4a	5087	0.9710816	51	0.1810295	99.74509804	5
FBtr0079216	CG9140	7771	0.9843457	74	0.208516	105.0135135	3
FBtr0079217	slmo	2959	0.9681677	32	0.2088509	92.46875	0
FBtr0079218	slmo	2919	0.9923011	30	0.2044483	97.3	1
FBtr0079288	CoVb	3464	0.9561043	13	0.223594	266.4615385	2
FBtr0079445	CG5261	9196	0.9503836	260	0.513555	35.36923077	7
FBtr0079500	Acp1	2812	0.9587629	13	0.1890034	216.3076923	1
FBtr0079546	RpL36A	1703	0.9737418	17	0.2997812	100.1764706	2
FBtr0079565	Rack1	7357	0.9891501	25	0.198915	294.28	2
FBtr0079701	Bace	9452	0.9975	40	0.2025	236.3	0
FBtr0079713	Peritrophin-15a	430	0.9556786	7	0.1468144	61.42857143	1
FBtr0079724	RpS13	3204	0.974359	45	0.0705128	71.2	2
FBtr0079796	Ggamma30A	987	0.9793916	21	0.1511286	47	2
FBtr0079888	RpL13	1580	0.969697	21	0.1414141	75.23809524	3
FBtr0079905	RpS2	3156	0.9781659	30	0.1877729	105.2	1
FBtr0079946	RpL7	3968	0.9776286	40	0.3344519	99.2	2
FBtr0080016	RpS27A	4308	0.9932432	13	0.2179054	331.3846154	2
FBtr0080050	Mdh1	2248	0.9732528	38	0.2847282	59.15789474	1
FBtr0080127	CG17108	3194	0.9871134	52	0.1176976	61.42307692	0
FBtr0080167	l(2)06225	4007	1	54	0.2231237	74.2037037	1
FBtr0080182	porin	5235	0.9881495	84	0.1203282	62.32142857	3
FBtr0080242	CG31705	1877	0.9596808	76	0.1768165	24.69736842	2
FBtr0080306	CG6770	3798	1	352	0.4384106	10.78977273	0
FBtr0080418	Vha68-2	5312	0.9786806	54	0.1604988	98.37037037	4
FBtr0080419	Vha68-2	5114	0.9768669	51	0.1558442	100.2745098	4
FBtr0080524	RpL24	2929	0.9926199	10	0.1156212	292.9	2
FBtr0080542	b	1469	0.9545903	35	0.171767	41.97142857	2
FBtr0080603	CG15293	1436	0.9656616	36	0.1432161	39.88888889	1
FBtr0080771	l(2)35Di	1837	0.9766764	3	0.0218659	612.3333333	2
FBtr0080889	Cyt-c-p	7633	0.9632893	24	0.2980911	318.0416667	1
FBtr0080895	Mhc	88213	0.9823036	177	0.1843131	498.3785311	17
FBtr0080896	Mhc	89458	0.9896521	181	0.1889622	494.2430939	17

Figure S3.1 (Continued)

FBtr0080897	Mhc	91781	0.9977505	175	0.1894121	524.4628571	17
FBtr0080898	Mhc	88274	0.9865027	179	0.1888122	493.150838	17
FBtr0080899	Mhc	87217	0.9815537	179	0.1865627	487.2458101	17
FBtr0080900	Mhc	89519	0.9938512	183	0.1934613	489.1748634	17
FBtr0080901	Mhc	89409	0.9890522	181	0.1889622	493.9723757	17
FBtr0080902	Mhc	88462	0.9889022	183	0.1912118	483.3989071	17
FBtr0080903	Mhc	88523	0.9931014	185	0.1957109	478.5027027	17
FBtr0080905	Mhc	95833	0.9974906	194	0.193364	493.9845361	18
FBtr0080906	Mhc	96643	0.9979079	202	0.1910739	478.4306931	18
FBtr0080907	Mhc	94532	0.9979079	194	0.1910739	487.2783505	18
FBtr0081030	Arr1	2742	0.9902439	34	0.138676	80.64705882	3
FBtr0081089	RpS26	4286	0.959375	15	0.1640625	285.7333333	1
FBtr0081090	RpS26	4218	0.9595646	15	0.163297	281.2	1
FBtr0081122	CG10570	1477	0.9620563	24	0.2031824	61.54166667	1
FBtr0081253	ref(2)P	2178	0.9615225	107	0.3582954	20.35514019	2
FBtr0081300	ColV	7858	0.9890561	10	0.1450068	785.8	2
FBtr0081301	ColV	7739	0.9874372	10	0.1331658	773.9	2
FBtr0081473	Acon	35918	0.9960245	75	0.1420311	478.9066667	3
FBtr0081639	alphaTub84B	3898	0.9942923	105	0.239726	37.12380952	1
FBtr0081855	CG9603	1026	0.9693053	18	0.1550889	57	2
FBtr0081882	CG8036	2331	0.957047	59	0.1583893	39.50847458	2
FBtr0081920	CG8369	734	1	24	0.4524362	30.58333333	2
FBtr0081930	CG9836	2636	0.9675393	79	0.5078534	33.36708861	2
FBtr0081978	VhaM8.9	1612	0.9856948	55	0.1968665	29.30909091	2
FBtr0082103	Crc	1515	0.9853147	55	0.3370629	27.54545455	3
FBtr0082136	RpS29	2566	0.9820717	7	0.2131474	366.5714286	2
FBtr0082158	MtnA	1956	0.9817629	11	0.3434651	177.8181818	1
FBtr0082344	Tctp	3538	0.9893823	127	0.5463321	27.85826772	0
FBtr0082346	RpL3	7376	0.9985518	96	0.2273715	76.83333333	5
FBtr0082358	CG5214	6478	0.9825803	55	0.1431682	117.7818182	6
FBtr0082370	RpS25	3064	0.9864078	4	0.1417476	766	2
FBtr0082372	SdhC	1814	0.9756098	37	0.3597561	49.02702703	1
FBtr0082474	CoVa	6551	0.9714693	16	0.2524964	409.4375	0
FBtr0082534	Lk6	8956	0.9625113	138	0.1458898	64.89855072	5
FBtr0082535	Lk6	7368	0.9640613	137	0.1437547	53.7810219	5
FBtr0082598	Cyp9f2	2759	0.9664391	24	0.1353811	114.9583333	3
FBtr0082607	GstD1	1179	0.9805353	63	0.3990268	18.71428571	1
FBtr0082626	desat1	5441	0.9789636	56	0.1682914	97.16071429	4
FBtr0082627	desat1	5924	0.9782071	62	0.1817732	95.5483871	4
FBtr0082628	desat1	5436	0.9596115	55	0.1610429	98.83636364	4
FBtr0082629	desat1	5436	0.9539939	57	0.1673408	95.36842105	4
FBtr0082630	desat1	5454	0.9839711	55	0.1628749	99.16363636	4
FBtr0082670	Vha55	6063	0.9628949	107	0.1667891	56.6635514	3
FBtr0082671	Vha55	5863	0.9714071	108	0.1745673	54.28703704	2
FBtr0082931	CG3321	2347	0.9981584	47	0.3333333	49.93617021	1
FBtr0082962	His4r	1416	0.9618056	11	0.1475694	128.7272727	2
FBtr0083030	Mf	6715	0.9901599	17	0.196802	395	5
FBtr0083032	Mf	4618	0.9550971	34	0.3009709	135.8235294	4
FBtr0083035	GlyS	5721	0.977885	57	0.2054684	100.3684211	4
FBtr0083037	GlyS	5577	0.9705882	58	0.2023994	96.15517241	4
FBtr0083055	Hsc70-4	6370	0.9798599	111	0.2535026	57.38738739	1
FBtr0083056	Hsc70-4	6381	0.9826198	111	0.245443	57.48648649	0
FBtr0083057	Hsc70-4	7552	0.9923986	114	0.2592905	66.24561404	1
FBtr0083058	Hsc70-4	6370	0.9815789	111	0.2539474	57.38738739	1
FBtr0083059	Hsc70-4	6376	0.952381	118	0.2589286	54.03389831	1
FBtr0083060	Hsc70-4	6373	0.9781958	111	0.2475417	57.41441441	1
FBtr0083063	Oscp	9690	0.9987531	12	0.1658354	807.5	2
FBtr0083078	Tm2	5949	0.978013	41	0.1807818	145.097561	3
FBtr0083143	Act88F	23367	0.9981191	28	0.1510972	834.5357143	2
FBtr0083154	CG18522	5049	0.9694891	147	0.1289026	34.34693878	5
FBtr0083164	CG5399	3018	0.9514964	18	0.1960784	167.6666667	2
FBtr0083191	ND23	2242	0.995116	23	0.2918193	97.47826087	2
FBtr0083563	Mdh2	9976	0.988764	59	0.3116105	169.0847458	3
FBtr0083712	NP15.6	2517	0.9909639	80	0.561747	31.4625	0
FBtr0083727	ATPsyn-d	3960	0.9717115	40	0.2446959	99	1
FBtr0083728	ATPsyn-d	4816	0.9919137	39	0.212938	123.4871795	1
FBtr0083804	Vha13	1290	0.9743955	124	0.1337127	10.40322581	2

Figure S3.1 (Continued)

FBtr0083857	ninaE	8376	0.9874765	42	0.1540388	199.4285714	4
FBtr0083890	MtnB	466	0.978125	2	0.1	233	1
FBtr0083935	CG4000	1583	0.989	154	0.685	10.27922078	2
FBtr0083964	RpS20	2783	0.964539	9	0.212766	309.2222222	3
FBtr0083969	RpS30	2894	0.9923518	10	0.1357553	289.4	2
FBtr0083970	RpS30	2865	0.9881657	10	0.1400394	286.5	1
FBtr0083991	CG17273	1438	0.9523321	63	0.2004101	22.82539683	4
FBtr0084153	fit	1220	0.9585688	13	0.2278719	93.84615385	0
FBtr0084172	ND42	3917	0.9905729	32	0.173314	122.40625	2
FBtr0084190	CG6439	7361	0.9903278	47	0.2294465	156.6170213	5
FBtr0084213	PyK	15855	0.9878665	57	0.1501517	278.1578947	3
FBtr0084214	PyK	16668	0.995396	37	0.160221	450.4864865	3
FBtr0084255	Pebp1	1468	0.9643963	24	0.3854489	61.16666667	0
FBtr0084410	RpS3	4118	0.9634956	21	0.2488938	196.0952381	1
FBtr0084432	ATPsyn-Cf6	3142	0.9923955	12	0.256654	261.8333333	2
FBtr0084466	CG10219	2083	0.9709172	48	0.3948546	43.39583333	3
FBtr0084847	tobi	1990	0.9724334	44	0.1074144	45.22727273	1
FBtr0084879	CG5107	1540	0.9590288	17	0.1487102	90.58823529	1
FBtr0084892	Npl4	8261	0.9800724	37	0.1644022	223.2702703	7
FBtr0084893	Npl4	8401	0.9932261	99	0.1943268	84.85858586	7
FBtr0084901	CG5028	6088	0.9808102	38	0.1911869	160.2105263	5
FBtr0084932	RpL27	3501	0.9944134	34	0.4283054	102.9705882	1
FBtr0084994	Ald	27217	0.9946889	98	0.2207891	277.7244898	3
FBtr0084995	Ald	18622	0.9984301	18	0.1562009	1034.555556	3
FBtr0085001	Ald	15594	0.9535947	46	0.2013072	339	3
FBtr0085077	CG6295	8517	1	44	0.3098075	193.5681818	1
FBtr0085094	BM-40-SPARC	1356	0.9702537	103	0.2992126	13.16504854	2
FBtr0085153	CG17192	3699	0.9838403	16	0.1568441	231.1875	1
FBtr0085195	Mlc1	5957	0.9920635	27	0.2414966	220.6296296	4
FBtr0085196	Mlc1	5999	0.9935414	27	0.2292788	222.1851852	5
FBtr0085248	RpL4	5149	0.9914286	42	0.1385714	122.5952381	3
FBtr0085366	CG11876	6207	0.9725086	37	0.1250859	167.7567568	5
FBtr0085369	CG11876	6060	0.994228	34	0.1183261	178.2352941	4
FBtr0085384	Pglym78	4101	0.971831	56	0.3961267	73.23214286	2
FBtr0085392	Ef1gamma	4000	0.9804849	127	0.3264341	31.49606299	3
FBtr0085393	Ef1gamma	3992	0.9778684	126	0.3127548	31.68253968	2
FBtr0085463	Obp99c	1411	0.9982609	24	0.4295652	58.79166667	1
FBtr0085502	Jon99Ciii	9600	0.995221	47	0.3990442	204.2553191	0
FBtr0085511	Jon99Ci	1969	0.9735391	14	0.1532525	140.6428571	1
FBtr0085512	Jon99Cii	9934	0.9883991	47	0.387471	211.3617021	0
FBtr0085535	CG7834	1112	0.9671814	30	0.2905405	37.06666667	2
FBtr0085536	CG7834	1090	0.9929221	27	0.2780586	40.37037037	1
FBtr0085539	ATPsyn-gamma	8442	0.9935185	24	0.1935185	351.75	1
FBtr0085541	ATPsyn-gamma	9289	0.9939966	26	0.1921098	357.2692308	0
FBtr0085583	Tpi	3056	0.9820144	38	0.1906475	80.42105263	2
FBtr0085592	RpL32	2588	0.9961832	8	0.1908397	323.5	2
FBtr0085603	CG7920	7356	0.9513089	40	0.2643979	183.9	3
FBtr0085632	Fer1HCH	5983	0.9920705	33	0.2986784	181.3030303	3
FBtr0085633	Fer1HCH	6337	0.9926651	37	0.2974735	171.2702703	3
FBtr0085634	Fer1HCH	6331	0.9924179	33	0.2855939	191.8484848	3
FBtr0085635	Fer1HCH	5493	0.9607351	40	0.3116124	137.325	2
FBtr0085713	Sap-r	7726	0.9927326	72	0.1241279	107.3055556	6
FBtr0085714	Sap-r	7377	0.9725291	58	0.1125709	127.1896552	5
FBtr0085745	l(3)03670	1773	0.9851064	47	0.1361702	37.72340426	2
FBtr0085772	CG1746	32177	0.9728978	36	0.2154274	893.8055556	3
FBtr0085774	CG1746	31139	0.9977728	36	0.2301411	864.9722222	2
FBtr0085800	CycG	4916	0.9869281	46	0.1877996	106.8695652	5
FBtr0085803	CycG	5110	0.9549945	54	0.1822173	94.62962963	5
FBtr0085804	RpL6	4310	0.9846154	28	0.132967	153.9285714	2
FBtr0085805	RpL6	5175	0.9865564	29	0.1416753	178.4482759	1
FBtr0085864	#N/A	2842	0.9716088	53	0.3911672	53.62264151	#N/A
FBtr0085892	His1:CG31617	2467	0.9933185	133	0.6636971	18.54887218	0
FBtr0085893	His2A:CG31618	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0085896	Lamp1	1871	0.9545168	40	0.1718256	46.775	3
FBtr0085911	Ef2b	18697	0.9920606	100	0.1613136	186.97	4
FBtr0085961	RpL21	3681	0.9723127	13	0.237785	283.1538462	1
FBtr0086150	Vha16-1	14292	0.9911754	128	0.2499118	111.65625	3

Figure S3.1 (Continued)

FBtr0086151	Vha16-1	14366	0.9907076	117	0.2323088	122.7863248	3
FBtr0086152	Vha16-1	14392	0.9661188	122	0.2358269	117.9672131	2
FBtr0086153	Vha16-1	14367	0.9861506	117	0.2308239	122.7948718	3
FBtr0086156	SdhB	3794	0.9835931	159	0.3281378	23.86163522	2
FBtr0086216	CG18067	1515	0.9913043	9	0.1428571	168.3333333	1
FBtr0086273	RpS18	2644	0.9771615	11	0.2120718	240.3636364	3
FBtr0086303	CG9090	9676	0.9928161	64	0.2104885	151.1875	3
FBtr0086477	Obp56d	658	0.9630873	10	0.135906	65.8	1
FBtr0086533	RpL11	4012	0.9798271	16	0.2463977	250.75	4
FBtr0086536	betaTub56D	1692	0.9795412	33	0.1680099	51.27272727	1
FBtr0086552	SdhA	12473	0.9805195	158	0.2456073	78.94303797	4
FBtr0086553	SdhA	12548	0.973849	162	0.2526703	77.45679012	3
FBtr0086554	SdhA	12471	0.9832317	156	0.238186	79.94230769	4
FBtr0086585	CG10737	3732	0.9538513	83	0.1503984	44.96385542	14
FBtr0086701	Pepck	4137	0.9574126	49	0.1361993	84.42857143	1
FBtr0086727	CG15068	428	0.973545	9	0.2380952	47.55555556	1
FBtr0086903	CG6484	1867	0.9703844	54	0.2173913	34.57407407	1
FBtr0086905	CG14482	974	1	34	0.5910653	28.64705882	1
FBtr0086985	CG11400	1176	1	9	0.1364942	130.6666667	0
FBtr0087004	Amy-p	34555	0.9993742	131	0.3667084	263.778626	0
FBtr0087005	GstS1	8792	0.9935806	65	0.2232525	135.2615385	4
FBtr0087006	GstS1	9478	0.9920635	87	0.2261905	108.9425287	4
FBtr0087105	RpLP2	2738	0.9732888	22	0.3722872	124.4545455	0
FBtr0087307	Gpo-1	9572	0.9871795	64	0.188172	149.5625	7
FBtr0087308	Gpo-1	10454	0.9895708	67	0.1945447	156.0298507	7
FBtr0087309	Gpo-1	10178	0.9894934	97	0.2213884	104.9278351	7
FBtr0087335	Vha36-1	1569	0.9713494	75	0.2939002	20.92	0
FBtr0087440	CG12859	757	0.9790795	17	0.3535565	44.52941176	1
FBtr0087494	CG30197	442	0.98125	9	0.2020833	49.11111111	3
FBtr0087560	Arc1	4322	0.9940878	137	0.1870777	31.54744526	0
FBtr0087591	#N/A	3885	0.9725752	39	0.2160535	99.61538462	#N/A
FBtr0087592	Cp1	3970	0.9793609	39	0.2150466	101.7948718	3
FBtr0087654	IM10	2277	0.9637362	170	0.7208791	13.39411765	3
FBtr0087656	CG33470	2548	0.9663503	224	0.7385677	11.375	3
FBtr0087732	AGBE	2721	0.9741379	143	0.2228448	19.02797203	3
FBtr0087746	CG4716	1325	0.9795918	25	0.3437991	53	1
FBtr0087747	CG4716	1607	0.9768519	30	0.3217593	53.56666667	0
FBtr0087783	bic	1235	0.9837728	57	0.1673428	21.66666667	0
FBtr0087796	CG13324	2464	0.9958071	13	0.197065	189.5384615	0
FBtr0087854	CG12374	28681	0.9986348	33	0.2006826	869.1212121	4
FBtr0087861	ox	2459	0.9749216	96	0.2664577	25.61458333	1
FBtr0088013	Oda	8409	0.9626007	115	0.3124281	73.12173913	2
FBtr0088014	Oda	8495	0.9710445	124	0.3293692	68.50806452	2
FBtr0088035	Ef1alpha48D	16194	0.9980286	43	0.1971415	376.6046512	1
FBtr0088122	betaTry	14113	1	34	0.2506234	415.0882353	0
FBtr0088123	CG30031	7960	0.9864365	28	0.2552404	284.2857143	0
FBtr0088124	deltaTry	7948	0.9975309	28	0.2555556	283.8571429	0
FBtr0088158	CG30025	8061	1	28	0.2555556	287.8928571	0
FBtr0088159	gammaTry	7960	0.9876543	28	0.2555556	284.2857143	0
FBtr0088160	epsilonTry	4628	0.9903615	100	0.5795181	46.28	0
FBtr0088161	alphaTry	15415	0.9872093	57	0.3476744	270.4385965	0
FBtr0088397	CoVIIc	2544	0.9928741	16	0.1995249	159	2
FBtr0088413	14-3-3zeta	2981	0.9566457	18	0.1158015	165.6111111	6
FBtr0088421	Pfk	6439	0.9611042	34	0.1282936	189.3823529	7
FBtr0088422	Pfk	7143	0.9723489	42	0.1342645	170.0714286	7
FBtr0088525	RpL31	2324	0.96139	19	0.2316602	122.3157895	2
FBtr0088527	RpL31	2133	0.9632653	16	0.2142857	133.3125	2
FBtr0088587	VhaAC45	2521	0.9597574	27	0.1576626	93.37037037	4
FBtr0088679	Pgi	10047	0.9562624	102	0.245328	98.5	4
FBtr0088709	PGRP-SC2	602	0.9513513	19	0.3873874	31.68421053	0
FBtr0088759	Mal-A1	5869	0.9956141	61	0.2001096	96.21311475	2
FBtr0088816	Obp44a	1180	0.9984472	75	0.5854037	15.73333333	1
FBtr0088872	ACC	7852	0.9529412	320	0.2840154	24.5375	12
FBtr0088898	cathD	1569	0.9577364	49	0.3517192	32.02040816	1
FBtr0089055	Cyp9b2	1256	0.95021	35	0.1391722	35.88571429	3
FBtr0089105	CG1970	4745	0.9802924	38	0.2403052	124.8684211	5
FBtr0089175	RpS3A	6772	0.9934066	22	0.2021978	307.8181818	1

Figure S3.1 (Continued)

FBtr0089176	RpS3A	6196	0.9929742	23	0.234192	269.3913043	2
FBtr0089186	ATPsyn-beta	32325	0.9976304	42	0.2233412	769.6428571	2
FBtr0089187	ATPsyn-beta	32131	0.9827685	45	0.2257323	714.0222222	1
FBtr0089188	Rfabg	51078	0.9973753	426	0.2149315	119.9014085	7
FBtr0089324	Lsp2	5761	0.9973958	76	0.2829861	75.80263158	0
FBtr0089329	Inos	2042	0.9738167	28	0.1540785	72.92857143	4
FBtr0089422	RpS7	4606	0.9835841	19	0.24487	242.4210526	3
FBtr0089497	Gdh	5822	0.9833024	238	0.3167903	24.46218487	4
FBtr0089498	Gdh	5855	0.9835991	241	0.3248292	24.29460581	5
FBtr0089510	Atpalpha	12121	0.9666115	113	0.200883	107.2654867	9
FBtr0089511	Atpalpha	13940	0.9541762	142	0.1972506	98.16901408	8
FBtr0089516	Atpalpha	13827	0.9508436	144	0.1980837	96.02083333	8
FBtr0089517	AnxB9	1409	0.9664537	56	0.2036741	25.16071429	4
FBtr0089518	AnxB9	1482	0.9755068	117	0.2407095	12.66666667	4
FBtr0089562	Zasp66	2115	0.9587459	16	0.1443894	132.1875	7
FBtr0089563	Zasp66	3921	0.9883314	42	0.1616103	93.35714286	8
FBtr0089566	Zasp66	1958	0.9526316	21	0.1421053	93.23809524	6
FBtr0089568	Zasp66	3365	0.9844804	38	0.159919	88.55263158	6
FBtr0089630	CG10910	3621	0.959204	16	0.0661692	226.3125	2
FBtr0089746	#N/A	29763	0.991453	102	0.2064777	291.7941176	#N/A
FBtr0089747	Mlc2	34396	0.9916885	101	0.1942257	340.5544554	2
FBtr0089766	Cyp6d5	3138	0.9798781	26	0.1353659	120.6923077	4
FBtr0089793	bt	73623	0.9779481	632	0.1362561	116.4920886	40
FBtr0089959	Tm1	7786	0.97981	28	0.1538005	278.0714286	9
FBtr0089967	Tm1	7028	0.9855769	23	0.1105769	305.5652174	9
FBtr0091464	#N/A	35778	0.9879518	584	0.1843691	61.26369863	#N/A
FBtr0091805	His1:CG33801	2431	0.9755011	135	0.6614699	18.00740741	0
FBtr0091808	His1:CG33804	2442	0.9888641	131	0.6536748	18.64122137	0
FBtr0091811	His1:CG33807	2420	0.9832962	136	0.6815145	17.79411765	0
FBtr0091812	His2A:CG33808	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0091814	His1:CG33810	2451	0.9933185	134	0.6837416	18.29104478	0
FBtr0091817	His1:CG33813	2467	0.9933185	133	0.6636971	18.54887218	0
FBtr0091818	His2A:CG33814	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0091820	His1:CG33816	2471	0.9933185	133	0.6636971	18.57894737	0
FBtr0091821	His2A:CG33817	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0091823	His1:CG33819	2451	0.9933185	134	0.6837416	18.29104478	0
FBtr0091824	His2A:CG33820	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0091826	His1:CG33822	2451	0.9933185	134	0.6837416	18.29104478	0
FBtr0091827	His2A:CG33823	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0091829	His1:CG33825	2451	0.9933185	134	0.6837416	18.29104478	0
FBtr0091830	His2A:CG33826	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0091832	His1:CG33828	2451	0.9933185	134	0.6837416	18.29104478	0
FBtr0091833	His2A:CG33829	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0091835	His1:CG33831	2451	0.9933185	134	0.6837416	18.29104478	0
FBtr0091836	His2A:CG33832	3195	0.9525483	230	0.8242531	13.89130435	0
FBtr0091838	His1:CG33834	2194	0.986637	136	0.6815145	16.13235294	0
FBtr0091841	His1:CG33837	2469	0.9933185	135	0.6614699	18.28888889	0
FBtr0091844	His1:CG33840	2469	0.9933185	135	0.6614699	18.28888889	0
FBtr0091847	His1:CG33843	2469	0.9933185	135	0.6614699	18.28888889	0
FBtr0091850	His1:CG33846	2469	0.9933185	135	0.6614699	18.28888889	0
FBtr0091853	His1:CG33849	2469	0.9933185	135	0.6614699	18.28888889	0
FBtr0091856	His1:CG33852	2449	0.9933185	136	0.6815145	18.00735294	0
FBtr0091859	His1:CG33855	2283	0.9743875	134	0.6837416	17.03731343	0
FBtr0091862	His1:CG33858	2283	0.9743875	134	0.6837416	17.03731343	0
FBtr0091865	His1:CG33861	2303	0.9743875	133	0.6636971	17.31578947	0
FBtr0091868	His1:CG33864	2469	0.9933185	135	0.6614699	18.28888889	0
FBtr0100164	RpL13	1580	0.9922481	21	0.1447028	75.23809524	3
FBtr0100182	14-3-3zeta	1929	0.9691992	11	0.1273101	175.3636364	6
FBtr0100183	14-3-3zeta	3064	0.9629207	18	0.1158015	170.2222222	6
FBtr0100231	RpL41	2926	1	10	0.3450479	292.6	2
FBtr0100289	ref(2)P	2126	0.9725596	104	0.3738192	20.44230769	2
FBtr0100321	fabp	2700	0.9635854	16	0.1694678	168.75	2
FBtr0100379	CG3164	3007	0.9521648	57	0.1714385	52.75438596	6
FBtr0100387	Zasp52	5646	0.9663098	107	0.1807305	52.76635514	14
FBtr0100388	Zasp52	15206	0.9735778	264	0.2280938	57.59848485	14
FBtr0100432	Jon25Bi	7481	0.9861432	31	0.2898383	241.3225806	0
FBtr0100479	Gapdh1	17086	1	143	0.4038911	119.4825175	0

Figure S3.1 (Continued)

FBtr0100482	Pglym78	4072	0.9823875	52	0.3806262	78.30769231	2
FBtr0100483	Pglym78	4060	0.9890329	51	0.3668993	79.60784314	2
FBtr0100485	GlyP	14067	0.9958492	91	0.2161882	154.5824176	3
FBtr0100541	RpS13	3204	1	45	0.0723684	71.2	2
FBtr0100561	up	12347	0.9876118	16	0.1266346	771.6875	7
FBtr0100563	up	12512	0.9881501	17	0.1310072	736	9
FBtr0100589	Adh	22148	1	19	0.1596639	1165.684211	3
FBtr0100590	Adh	20800	0.9990157	20	0.1584646	1040	2
FBtr0100594	Adh	22100	0.9975904	49	0.1726908	451.0204082	3
FBtr0100620	Prx5	1217	0.9986358	28	0.2755798	43.46428571	2
FBtr0100662	Act5C	8527	0.964032	44	0.19627	193.7954545	1
FBtr0100663	Act5C	8437	0.9756224	39	0.1950208	216.3333333	1
FBtr0100857	mt:ND2	11253	0.9941521	904	0.837232	12.44800885	0
FBtr0100861	mt:CoI	136896	0.9993489	928	0.4726562	147.5172414	0
FBtr0100863	mt:CoII	39466	1	65	0.3886463	607.1692308	0
FBtr0100866	mt:ATPase8	984	1	6	0.2777778	164	0
FBtr0100867	mt:ATPase6	33448	1	24	0.2918518	1393.666667	0
FBtr0100868	mt:CoIII	46090	1	47	0.4195184	980.6382979	0
FBtr0100870	mt:ND3	5115	1	11	0.3474576	465	0
FBtr0100877	mt:ND5	51780	0.9994203	85	0.2243478	609.1764706	0
FBtr0100879	mt:ND4	30649	0.9932886	68	0.2818792	450.7205882	0
FBtr0100880	mt:ND4L	3569	1	355	0.7491409	10.05352113	0
FBtr0100884	mt:Cyt-b	50188	0.9991205	639	0.4564644	78.54147105	0
FBtr0100886	mt:ND1	22282	1	52	0.371672	428.5	0
FBtr0110844	noe	4150	0.9800664	28	0.166113	148.2142857	0
FBtr0110845	noe	4436	0.9790419	30	0.1656687	147.8666667	0
FBtr0110872	l(2)06225	4453	0.95	52	0.19	85.63461538	2
FBtr0111048	Zasp52	5890	0.9694809	120	0.1782658	49.08333333	15
FBtr0111120	RpL38	2441	0.9850374	3	0.1022444	813.6666667	1
FBtr0111129	RpL5	6389	0.9592215	11	0.1167748	580.8181818	3
FBtr0111132	RpL5	6397	0.9863281	10	0.1064453	639.7	4
FBtr0112363	CG34172	957	0.9932886	3	0.1744967	319	2
FBtr0112413	CG34220	14777	1	202	0.5171952	73.15346535	1
FBtr0112526	CG34324	3965	0.9827236	84	0.324187	47.20238095	1
FBtr0112532	CG34330	491	0.9677419	16	0.2225806	30.6875	0
FBtr0112791	Argk	13323	0.9530745	38	0.2195254	350.6052632	2
FBtr0112860	Nc73EF	17011	0.9851043	157	0.1730387	108.3503185	11
FBtr0113101	CG10320	1952	0.9776675	5	0.1240695	390.4	2
FBtr0113140	DOR	4168	0.95383	138	0.1448059	30.20289855	4
FBtr0113265	CG5778	1245	0.9599156	33	0.3565401	37.72727273	2
FBtr0113290	CG5028	6856	0.9962686	73	0.2288557	93.91780822	5
FBtr0113360	CG30118	2433	0.9554962	82	0.3119715	29.67073171	9
FBtr0113464	Unc-89	36514	0.9902313	577	0.1791966	63.28249567	37
FBtr0113742	RpL15	5777	0.9915374	7	0.1184767	825.2857143	2
FBtr0114472	#N/A	3013	0.9590214	81	0.2629969	37.19753086	#N/A
FBtr0114473	#N/A	1654	0.9529873	31	0.2272282	53.35483871	#N/A
FBtr0114536	Pdh	2838	0.998913	16	0.125	177.375	2
FBtr0114537	Pdh	2806	0.9967141	16	0.1259584	175.375	2
FBtr0114548	Idh	4279	0.9634009	24	0.1345721	178.2916667	2
FBtr0273322	Mal-A6	3626	0.9861702	42	0.1718085	86.33333333	2
FBtr0273393	CG3214	2640	0.9644013	22	0.1666667	120	4
FBtr0273398	pst	3592	0.9533516	98	0.24853	36.65306122	6
FBtr0273399	pst	3571	0.9725696	87	0.2303348	41.04597701	5
FBtr0290272	CG7203	2096	0.9729345	37	0.3005698	56.64864865	1
FBtr0290312	CG10737	3753	0.9553366	63	0.1484401	59.57142857	14
FBtr0290316	CG10737	3285	0.9543164	63	0.1518307	52.14285714	14
FBtr0299696	Mlp60A	934	0.98627	8	0.173913	116.75	2
FBtr0299697	Mlp60A	4509	0.9961109	34	0.1696646	132.6176471	8
FBtr0299869	RpL10	3788	0.9884319	15	0.1928021	252.5333333	4
FBtr0299870	RpL10	3674	0.9885932	16	0.2091255	229.625	4
FBtr0300283	CG7461	2304	0.9732938	115	0.2799209	20.03478261	0
FBtr0300395	CG9485	5018	0.9536751	257	0.2388306	19.52529183	7
FBtr0300425	Acon	36665	0.9986154	94	0.1481481	390.0531915	3
FBtr0300635	CG42502	1477	0.9620563	24	0.2031824	61.54166667	1
FBtr0300680	CG10320	1956	0.9779412	5	0.122549	391.2	2
FBtr0300730	Npl4	8195	0.9797048	34	0.1526753	241.0294118	7
FBtr0300828	RpS15Aa	4154	0.9719764	18	0.2920354	230.7777778	2

Figure S3.1 (Continued)

FBtr0300899	CG18624	1383	0.9854015	4	0.1094891	345.75	1
FBtr0301036	NDUFA8	1804	0.9511834	35	0.1967456	51.54285714	4
FBtr0301154	Zasp66	3764	0.9853837	47	0.1607795	80.08510638	7
FBtr0301156	Zasp66	3484	0.98713	39	0.1570142	89.33333333	7
FBtr0301340	bt	74280	0.9745942	648	0.1384218	114.6296296	45
FBtr0301561	#N/A	6065	0.9718593	10	0.1095477	606.5	#N/A
FBtr0301661	Vha55	5472	0.9916992	95	0.168457	57.6	3
FBtr0301708	Men-b	2952	0.9531981	48	0.124025	61.5	6
FBtr0301783	Ucrh	1369	0.9903846	4	0.1706731	342.25	2
FBtr0301784	Ucrh	1311	0.9700461	4	0.1635945	327.75	2
FBtr0301827	Mhc	90292	0.9893095	183	0.1896065	493.3989071	18
FBtr0301828	Mhc	88034	0.9663398	178	0.1820577	494.5730337	17
FBtr0301829	Mhc	92615	0.9977719	175	0.1876114	529.2285714	18
FBtr0301919	up	12218	0.9876118	10	0.1114935	1221.8	7
FBtr0301920	up	12220	0.9881735	11	0.1162943	1110.909091	9
FBtr0301921	up	12202	0.987984	10	0.1081442	1220.2	8
FBtr0301922	up	12533	0.9881501	24	0.1514154	522.2083333	10
FBtr0301923	up	12503	0.9879599	23	0.1438127	543.6086957	9
FBtr0301959	Tm1	7082	0.9864326	24	0.1237031	295.0833333	9
FBtr0301960	Tm1	5894	0.9864326	21	0.1125299	280.6666667	9
FBtr0302301	RpL28	2098	0.9711539	16	0.2692308	131.125	3
FBtr0302442	Ef1gamma	3984	0.979941	128	0.3345133	31.125	3
FBtr0302527	CG33346	2265	0.9762774	102	0.1322993	22.20588235	2
FBtr0302586	RpL3	7082	0.9567536	98	0.2049763	72.26530612	5
FBtr0302854	Phae2	1055	0.9593679	9	0.0541761	117.2222222	0
FBtr0303048	Actn	9280	0.9779189	170	0.2683001	54.58823529	9
FBtr0303096	Scsalpha	2564	0.9864	40	0.2208	64.1	3
FBtr0303099	CG33470	2276	0.9690266	170	0.7256637	13.38823529	3
FBtr0303859	#N/A	3416	0.9590588	100	0.2983529	34.16	#N/A
FBtr0303860	CG42837	3416	0.9563585	101	0.3054904	33.82178218	2
FBtr0304129	MtnE	613	0.9710982	12	0.2427746	51.08333333	1
FBtr0304693	RpS27	2466	0.9665211	15	0.1644833	164.4	2
FBtr0304812	CG43078	12957	0.9571664	402	0.2184043	32.23134328	4
FBtr0304813	CG43078	15347	0.9517544	431	0.2011773	35.60788863	8
FBtr0304814	CG43078	15347	0.9647667	419	0.1976912	36.62768496	6
FBtr0304815	CG43078	15370	0.9534187	428	0.199239	35.91121495	7
FBtr0305122	CG1746	29406	0.9688889	31	0.2340741	948.5806452	2
FBtr0305260	porin	6182	0.9645902	95	0.1416393	65.07368421	3
FBtr0305551	Vha68-2	4825	0.9648118	49	0.150982	98.46938776	4
FBtr0305669	RpL29	1059	0.990099	10	0.2409241	105.9	2
FBtr0305977	cp309	13928	0.9612514	195	0.1473106	71.42564103	10
FBtr0305979	cp309	13954	0.952391	203	0.1484095	68.73891626	12
FBtr0306086	CG4769	8884	1	22	0.2087336	403.8181818	5
FBtr0306237	CG1746	16621	0.9958449	18	0.1786704	923.3888889	2
FBtr0306630	Fhos	7439	0.9557669	198	0.163704	37.57070707	16
FBtr0306632	Fhos	9453	0.9732957	288	0.1778073	32.82291667	16
FBtr0306639	Prm	4364	0.9705015	51	0.1545723	85.56862745	2
FBtr0306657	Ald	15587	0.9534732	46	0.2018349	338.8478261	3
FBtr0307034	cype	2667	0.977387	7	0.2386935	381	2
FBtr0307035	cype	2567	0.9573991	10	0.3071749	256.7	2
FBtr0307492	Mhc	87996	0.9782768	169	0.1781888	520.6863905	17
FBtr0307493	Mhc	88737	0.9880078	173	0.1875281	512.9306358	17
FBtr0307494	Mhc	88656	0.9799297	170	0.1800469	521.5058824	17
FBtr0307495	Mhc	89942	0.9959221	185	0.1910588	486.172973	16
FBtr0307496	Mhc	90167	0.9898477	184	0.1892676	490.0380435	18
FBtr0307904	l(1)G0156	14167	0.9950715	168	0.4938393	84.32738095	4
FBtr0308034	DOR	4222	0.9752353	141	0.1515602	29.94326241	4
FBtr0308188	CG34172	942	0.9932203	3	0.1762712	314	2
FBtr0308192	RpS10b	2583	0.9712042	16	0.2159686	161.4375	1
FBtr0308233	up	12232	0.9887719	10	0.1136842	1223.2	7
FBtr0308234	up	12532	0.9864197	24	0.1425926	522.1666667	9
FBtr0308235	wupA	5744	0.9966997	265	0.1254125	21.6754717	9
FBtr0308236	wupA	4821	0.9727891	263	0.133139	18.33079848	8
FBtr0308237	wupA	4866	0.9865255	263	0.1318576	18.50190114	8
FBtr0308333	RpL15	5780	0.9911168	7	0.106599	825.7142857	2
FBtr0308334	RpL15	5683	0.9899135	7	0.1210375	811.8571429	2
FBtr0308595	Mpcp	4758	0.9599097	103	0.2964427	46.19417476	3

Figure S3.1 (Continued)

FBtr0308683	CoVIb	4486	0.9820359	30	0.3473054	149.5333333	1
FBtr0310086	Zasp52	6099	0.9608732	133	0.2035724	45.85714286	15
FBtr0310087	Zasp52	10009	0.9661795	184	0.2194155	54.39673913	13
FBtr0310088	Zasp52	9280	0.972396	168	0.2037956	55.23809524	14
FBtr0310136	up	12370	0.9881735	24	0.151117	515.4166667	10
FBtr0310137	up	12358	0.9881501	18	0.1408822	686.5555556	9
FBtr0310138	up	12340	0.9879599	17	0.1331104	725.8823529	8
FBtr0310274	RpS8	3243	0.9862259	21	0.369146	154.4285714	3
FBtr0310496	#N/A	2718	0.9709454	143	0.2241977	19.00699301	#N/A
FBtr0310535	CG14125	6349	0.9860465	35	0.1383721	181.4	1
FBtr0310658	Actn	9154	0.9655684	208	0.2760463	44.00961538	9
FBtr0310659	Actn	9047	0.9632946	209	0.2820054	43.28708134	9
FBtr0310661	Ald	16541	0.9647059	91	0.2773994	181.7692308	3
FBtr0329909	Zasp52	2182	0.9572083	86	0.195619	25.37209302	9
FBtr0329911	Zasp52	14451	0.9783468	183	0.2181129	78.96721311	6
FBtr0329912	Zasp52	5456	0.9540306	112	0.2038641	48.71428571	14
FBtr0329913	Zasp52	5855	0.9568346	120	0.2086331	48.79166667	14
FBtr0329914	Zasp52	5351	0.9530298	102	0.1954105	52.46078431	13
FBtr0329915	Zasp52	5758	0.952748	111	0.2060992	51.87387387	13
FBtr0329916	Zasp52	6289	0.9710889	128	0.1837341	49.1328125	15
FBtr0330021	CG43733	987	0.9793916	21	0.1511286	47	2
FBtr0330025	Men	4000	0.9839255	85	0.1489002	47.05882353	2
FBtr0330403	RpL37a	1452	0.952381	12	0.1445578	121	2
FBtr0330640	CG11455	647	0.9592326	34	0.529976	19.02941176	2
FBtr0330682	RpS2	3053	0.9512987	26	0.1525974	117.4230769	1
FBtr0331361	RpL27	3498	0.9944853	32	0.3841912	109.3125	1
FBtr0331425	Gbs-76A	5299	0.9550056	86	0.1484814	61.61627907	2
FBtr0331557	CG4169	12681	0.9993494	85	0.2940794	149.1882353	3
FBtr0331564	CG6020	6703	0.9865139	22	0.1078894	304.6818182	3
FBtr0331650	CG34172	949	0.970297	3	0.1716172	316.3333333	2
FBtr0331810	CG7430	7566	0.9847609	37	0.1529164	204.4864865	3
FBtr0331864	mtacp1	2127	0.985115	29	0.1393775	73.34482759	4
FBtr0331936	CG7712	1750	0.9542484	12	0.2663399	145.8333333	2
FBtr0332029	porin	6352	0.9790916	90	0.1319394	70.57777778	3
FBtr0332168	Fer2LCH	4462	0.9945005	66	0.1594867	67.60606061	2
FBtr0332169	Fer2LCH	4412	0.9847943	66	0.1556351	66.84848485	3
FBtr0332370	nrv1	2459	0.9626168	89	0.2637591	27.62921348	3
FBtr0332500	Pif1A	6397	0.954645	201	0.1494653	31.82587065	13
FBtr0332503	Pif1B	6397	0.954645	201	0.1494653	31.82587065	13
FBtr0332526	CG5261	10910	0.9870067	269	0.5096247	40.55762082	6
FBtr0332527	CG5261	11054	0.9873001	269	0.4981185	41.0929368	6
FBtr0332528	CG5261	10980	0.9874243	271	0.5030275	40.51660517	6
FBtr0332529	CG5261	11156	0.9876993	271	0.4920273	41.16605166	6
FBtr0332597	Aldh	7141	0.9510526	104	0.2484211	68.66346154	3
FBtr0332618	Gapdh2	8216	0.9912854	75	0.4139434	109.5466667	1
FBtr0332651	CG18135	3822	0.9587961	51	0.1010391	74.94117647	0
FBtr0332683	sls	50787	0.9546723	1130	0.1481796	44.94424779	30
FBtr0332723	fln	7315	0.9592834	61	0.3501629	119.9180328	3
FBtr0332731	pst	3571	0.9767055	87	0.2375208	41.04597701	6
FBtr0332820	CG43897	9017	0.9590457	192	0.2357853	46.96354167	5
FBtr0332822	CG43897	9406	0.9753484	166	0.2340121	56.6626506	6
FBtr0332824	CG43897	9582	0.9662958	173	0.2392016	55.38728324	6
FBtr0332825	CG43897	9408	0.9753748	166	0.2337616	56.6746988	6
FBtr0332826	CG43897	10405	0.9850863	138	0.2103611	75.39855072	4
FBtr0332832	CG43897	12530	0.9537507	218	0.2408347	57.47706422	5
FBtr0332833	CG43897	10114	0.9557158	190	0.2468246	53.23157895	6
FBtr0332834	CG43897	11595	0.9900955	150	0.2164839	77.3	4
FBtr0332835	CG43897	11717	0.9516524	187	0.2392901	62.65775401	5
FBtr0332947	Ef1beta	1683	0.9938195	19	0.1248455	88.57894737	1
FBtr0332966	Ef2b	18935	0.9930338	103	0.1609195	183.8349515	4
FBtr0333111	CG43078	15006	0.9502226	431	0.2015695	34.81670534	7
FBtr0333112	CG43078	15342	0.9578292	425	0.2022789	36.09882353	7
FBtr0333113	CG43078	14975	0.9597819	422	0.2041187	35.48578199	6
FBtr0333126	AcCoAS	2339	0.9634011	38	0.1393715	61.55263158	5
FBtr0333142	Ubi-p63E	33591	0.9809451	75	0.2747713	447.88	1
FBtr0333276	GlyS	5548	0.9729074	54	0.1985443	102.7407407	4
FBtr0333310	#N/A	7824	0.9811193	27	0.1402562	289.7777778	#N/A

Figure S3.1 (Continued)

FBtr0333311	#N/A	7836	0.9811321	27	0.1401617	290.2222222	#N/A
FBtr0333370	RpL18	2424	0.9695767	49	0.3796296	49.46938776	2
FBtr0333383	Est-6	2155	0.9569951	18	0.0990746	119.7222222	1
FBtr0333551	GstS1	8702	0.9819089	73	0.2295696	119.2054795	4
FBtr0333676	CG1970	4728	0.9665404	42	0.2493687	112.5714286	6
FBtr0333709	RpL26	3106	0.9601227	22	0.2269939	141.1818182	1
FBtr0333710	CG3819	2096	0.9528832	30	0.1462729	69.86666667	1
FBtr0333777	ATPsyn-Cf6	3306	0.9522184	15	0.2849829	220.4	2
FBtr0333801	Actn	9353	0.96843	225	0.2753128	41.56888889	9
FBtr0333802	Actn	9149	0.9655377	208	0.2762923	43.98557692	9
FBtr0333918	Tm1	6241	0.9864326	21	0.1117318	297.1904762	9
FBtr0333942	cp309	13923	0.9554935	200	0.1509353	69.615	12
FBtr0333943	cp309	13943	0.954643	195	0.1438988	71.5025641	11
FBtr0333944	cp309	13784	0.9500276	194	0.1424159	71.05154639	11
FBtr0334029	His4r	1449	0.954792	11	0.1537071	131.7272727	2
FBtr0334094	l(3)neo18	3770	0.9971989	16	0.2240896	235.625	3
FBtr0334327	skap	9376	0.9921094	208	0.1662283	45.07692308	6
FBtr0334328	skap	9372	0.9937304	258	0.1781609	36.3255814	7
FBtr0334482	bt	72917	0.9812534	604	0.1356031	120.7235099	34
FBtr0334483	bt	74339	0.9778902	637	0.1355186	116.7017268	43
FBtr0334627	Rfabg	51256	0.962839	486	0.2212047	105.4650206	7
FBtr0334706	CG7920	7351	0.9966499	40	0.2819654	183.775	3
FBtr0334842	RFeSP	4541	0.9874777	27	0.2576029	168.1851852	2
FBtr0334843	RFeSP	4456	0.989	22	0.221	202.5454545	2
FBtr0334890	CoVa	6495	1	16	0.2787402	405.9375	1
FBtr0335194	ND42	3896	0.9845397	32	0.167955	121.75	2
FBtr0335199	CG6455	3490	0.9580574	42	0.1449595	83.0952381	7
FBtr0335200	CG6455	3318	0.9753087	38	0.1393867	87.31578947	8
FBtr0335387	CG3523	15992	0.9769514	313	0.1746465	51.09265176	5
FBtr0335523	#N/A	24	0.972973	1	0.1351351	24	#N/A
FBtr0336613	Vha68-2	4836	0.9652385	56	0.172595	86.35714286	4
FBtr0336614	Vha68-2	5317	0.9723101	54	0.1578323	98.46296296	4
FBtr0336620	Ggamma30A	919	0.9764429	21	0.1813899	43.76190476	1
FBtr0336621	CG43733	919	0.9764429	21	0.1813899	43.76190476	1
FBtr0336648	Strn-Mlck	38968	0.9995338	208	0.1459207	187.3461538	7
FBtr0336668	l(1)G0156	11220	0.9921147	114	0.5290322	98.42105263	5
FBtr0336707	Gel	2761	0.9547389	48	0.1245648	57.52083333	6
FBtr0336895	wupA	5070	0.9844694	298	0.1691113	17.01342282	7
FBtr0336896	wupA	5081	0.9599359	298	0.1570513	17.05033557	8
FBtr0336897	wupA	5456	0.982925	300	0.1566444	18.18666667	8
FBtr0337053	Hsc70-4	6381	0.9811644	121	0.265839	52.73553719	1

We identified a set of transcripts featuring a high degree of read coverage on the sense strand. We designate transcripts with >= 95% sense strand coverage and a >=10:1 sense:antisense read count ratio as high exon coverage transcripts or "HECTs".

Figure S3.2. Differential expression analysis for HECTs

HECT	gene	logCPM	logFC_LTM	PValue_LTM	FDR_LTM	logFC_ODOR	PValue_ODOR	FDR_ODOR	logFC_SHOCK	PValue_SHOCI	FDR_SHOCK
FBtr0079701	Bace	10.32945118	3.620403329	5.10463E-21	3.97651E-18	2.627876557	5.48605E-09	4.27364E-06	4.256190358	4.80278E-20	3.74136E-17
FBtr0079056	Jon25Biii	8.030294733	3.255188346	1.14846E-15	4.47324E-13	0.730056526	0.141481456	0.999941333	2.860706811	2.27747E-09	1.36473E-07
FBtr0087854	CG12374	12.32874927	2.632599256	1.86599E-12	4.84535E-10	1.793674285	8.07969E-05	0.004495768	3.46124417	9.85298E-14	3.83774E-11
FBtr0084879	CG5107	7.574347614	2.770736306	5.7296E-12	1.11584E-09	1.872696981	7.58423E-05	0.004495768	2.516582747	7.14526E-08	3.34673E-06
FBtr0079713	Peritrophin-15a	5.564630877	3.184910136	9.76473E-12	1.52134E-09	-2.828751042	0.005930072	0.159293992	3.722063479	5.03955E-11	5.6083E-09
FBtr0088759	Mal-A1	9.576363078	2.49725059	1.92667E-11	2.50147E-09	1.30508617	0.003020341	0.084030208	1.778216004	4.00874E-05	0.001040936
FBtr0085153	CG17192	9.049134117	2.421718422	1.6136E-10	1.79571E-08	1.823846913	0.000117905	0.006123195	3.53598536	1.81256E-13	4.70662E-11
FBtr0088160	epsilonTry	9.074715894	2.3786969	2.06579E-10	2.01157E-08	1.514901585	0.000773528	0.027389927	3.008822871	3.34098E-11	4.33771E-09
FBtr0084255	Pebp1	6.569260228	2.783496261	4.15451E-10	3.59596E-08	1.359470141	0.008774301	0.207126692	3.317037426	9.68563E-12	1.88628E-09
FBtr0088161	alphaTry	11.37923134	2.200453451	1.84589E-09	1.43795E-07	1.495293128	0.000824308	0.027918957	2.328897845	2.48044E-07	1.01698E-05
FBtr0273322	Mal-A6	8.858642627	2.210382803	3.52842E-09	2.49876E-07	1.078762684	0.01670177	0.342386277	1.667997	0.000164	0.003193909
FBtr0085511	Jon99Ci	7.686652768	2.254264065	4.42556E-09	2.87292E-07	1.452602696	0.002501356	0.072168754	2.566291122	7.30353E-08	3.34673E-06
FBtr0085077	CG6295	10.34175254	2.154961193	5.65506E-09	3.38868E-07	1.429222196	0.001646783	0.049340137	2.713551657	3.45109E-09	1.92028E-07
FBtr0088159	gammaTry	10.37877016	2.077196184	1.32651E-08	6.40925E-07	1.558935995	0.0005304	0.020659085	2.219592464	1.00115E-06	3.39085E-05
FBtr0088123	CG30031	10.37877016	2.077196184	1.32651E-08	6.40925E-07	1.558935995	0.0005304	0.020659085	2.219592464	1.00115E-06	3.39085E-05
FBtr0088124	deltaTry	10.37777572	2.078544342	1.30416E-08	6.40925E-07	1.559365748	0.000527513	0.020659085	2.215535916	1.04827E-06	3.4025E-05
FBtr0088158	CG30025	10.39229208	2.073891997	1.39868E-08	6.40925E-07	1.543676239	0.000598689	0.022208494	2.202167473	1.21165E-06	3.77549E-05
FBtr0088122	betaTry	11.23906148	2.086481279	1.9047E-08	8.2431E-07	1.231449573	0.006658721	0.172904794	2.243423803	8.99562E-07	3.33695E-05
FBtr0100432	Jon25Bi	9.555398818	2.142621737	2.1031E-08	8.62273E-07	0.798803592	0.086051987	1	2.235062405	1.34979E-06	4.04419E-05
FBtr0079055	Jon25Bii	8.549135672	2.121878509	2.80383E-08	1.09209E-06	0.695959055	0.139273592	1	2.602572253	1.83901E-08	9.55057E-07
FBtr0085502	Jon99Ciii	10.73726917	1.932145563	2.05872E-07	7.63686E-06	1.173166125	0.010840075	0.241269102	2.818108392	1.31334E-09	9.92976E-08
FBtr0085512	Jon99Cii	10.77237955	1.923091035	2.31057E-07	8.18152E-06	1.173969413	0.01067131	0.241269102	2.805768841	1.52962E-09	9.92976E-08
FBtr0077040	Jon65Aiv	10.91616028	1.933304498	3.18892E-07	1.08007E-05	1.072585444	0.022321045	0.434702345	2.343115621	4.81161E-07	1.87412E-05
FBtr0084847	tobi	7.931481431	1.891585114	4.29184E-07	1.39306E-05	1.577051723	0.00100083	0.031185873	2.140044451	4.001E-06	0.000115436
FBtr0088709	PGRP-SC2	6.543769478	1.931778074	4.02038E-06	0.000125275	0.476463408	0.366910813	1	2.974932082	1.52108E-09	9.92976E-08
FBtr0086903	CG6484	7.778491615	1.759001276	5.73808E-06	0.000171922	0.249965303	0.598308307	1	1.77566859	8.06352E-05	0.001847495
FBtr0077041	Jon65Aiii	10.43650287	1.663028194	8.83111E-06	0.000254794	0.87032193	0.061064576	1	2.093795711	5.06365E-06	0.000140878
FBtr0083890	MtnB	6.043141561	2.051155162	1.13597E-05	0.000316042	-0.730190928	0.21782834	1	3.260458429	1.98409E-10	1.71734E-08
FBtr0077038	yip7	10.01245674	1.395532491	0.000173332	0.004656055	0.589596543	0.207885645	1	1.810158875	7.63079E-05	0.001801328
FBtr0083164	CG5399	8.631121752	1.384259124	0.000180857	0.004696244	0.395145988	0.392573563	1	1.716239777	0.000151494	0.003105621
FBtr0083154	CG18522	9.494026791	1.303176167	0.000250548	0.00629602	0.667675858	0.135180446	1	1.984479453	7.49882E-06	0.000201434
FBtr0112532	CG34330	6.020163889	1.472056411	0.000270501	0.006585005	0.266218186	0.623685063	1	1.556127745	0.002180709	0.027848726
FBtr0076389	CG18180	10.31882834	1.308776985	0.000448068	0.010577118	0.566555434	0.22861722	1	1.877131479	4.54348E-05	0.001106054
FBtr0087004	Amy-p	13.05950632	1.272139613	0.000727538	0.016669185	0.682910777	0.139192125	1	1.647468922	0.000369173	0.006390797
FBtr0087796	CG13324	8.342301808	1.223203793	0.001961573	0.043659011	0.077403556	0.875206117	1	1.098904698	0.026008872	0.082980761
FBtr0087560	Arc1	9.282015437	1.104206037	0.002288095	0.04951184	0.021216078	0.961626383	1	1.404052983	0.001161555	0.017072666
FBtr0089055	Cyp9b2	6.545634276	1.109458187	0.004225708	0.08662701	0.325423255	0.517957391	1	1.554253795	0.001288123	0.018582371
FBtr0112526	CG34324	9.043304547	1.008168282	0.004544997	0.090783396	-0.694504974	0.131340981	1	1.274904987	0.00425866	0.036059737
FBtr0333710	CG3819	7.859239329	1.115912627	0.00533871	0.101435494	0.565839357	0.258870484	1	1.528135503	0.001488994	0.020712974
FBtr0075048	CG3819	7.847583235	1.116425977	0.005322113	0.101435494	0.579179008	0.247833109	1	1.523884915	0.001531361	0.020928594
FBtr0075069	CG6839	11.27568614	1.013622693	0.007713066	0.143059015	0.355078492	0.448469104	1	1.269801618	0.006326514	0.04378475
FBtr0302854	Phae2	7.207076511	0.984864328	0.008547324	0.151326486	0.322287705	0.500460694	1	1.436502489	0.002117064	0.027848726
FBtr0336613	Vha68-2	9.106794464	0.930963537	0.009933009	0.168212807	0.088179555	0.84506902	1	1.163137873	0.00895607	0.049254962
FBtr0305551	Vha68-2	9.102613499	0.928261345	0.010148911	0.168212807	0.090586006	0.840893386	1	1.155680997	0.009401323	0.049254962
FBtr0075878	RpS12	8.411205496	0.907930168	0.01184735	0.182636759	0.022334229	0.960761551	1	1.473960406	0.000931891	0.013960451
FBtr0076633	RpL14	8.741912177	0.893166169	0.012581385	0.182636759	0.385245365	0.386386843	1	1.241227115	0.004884126	0.037684862
FBtr0114548	Idh	9.224094114	0.813626845	0.019572776	0.217817035	0.031204164	0.94555062	1	1.570161586	0.000438522	0.007268275
FBtr0086150	Vha16-1	11.04022178	0.828599993	0.01952864	0.217817035	0.171678667	0.695863746	1	1.185643082	0.006570929	0.044531937
FBtr0086153	Vha16-1	11.04043763	0.830036766	0.019290492	0.217817035	0.178917367	0.683763099	1	1.184324301	0.006644002	0.044531937
FBtr0086151	Vha16-1	11.04031351	0.829781695	0.019329058	0.217817035	0.178806336	0.683952611	1	1.184211391	0.006649917	0.044531937
FBtr0086152	Vha16-1	11.04183638	0.830132208	0.019286414	0.217817035	0.178904727	0.683787316	1	1.183377699	0.006688365	0.044531937
FBtr0076667	Idh	9.216929939	0.806834948	0.020626858	0.22011401	0.036139698	0.937008706	1	1.56609233	0.000455171	0.007315077
FBtr0076668	Idh	9.21590163	0.807940858	0.020457157	0.22011401	0.030552996	0.946685139	1	1.564785899	0.000460127	0.007315077
FBtr0332168	Fer2LCH	8.956971975	0.80793668	0.023231345	0.244556999	-0.098785913	0.825378754	1	1.648843783	0.000176371	0.003351042
FBtr0332169	Fer2LCH	8.940482159	0.806242014	0.023547886	0.244584039	-0.097900584	0.826965434	1	1.635841995	0.000198879	0.00368873
FBtr0082598	Cyp9f2	8.477754821	0.793937728	0.025546703	0.261853701	-0.188977801	0.678282028	1	1.17636306	0.008494598	0.049254962
FBtr0074522	wupA	12.75902557	-0.916608422	0.026355719	0.266637726	-2.178521708	2.6027E-05	0.0020275	-1.526213605	0.009457341	0.049254962
FBtr0087105	RpLP2	8.33060914	0.797024305	0.027330524	0.269690273	0.30527903	0.502691904	1	1.169028319	0.009359349	0.049254962
FBtr0336895	wupA	12.74721845	-0.913008776	0.028055415	0.273189601	-2.219842495	1.90447E-05	0.001984789	-1.554792496	0.008244331	0.049254962
FBtr0336896	wupA	12.74740543	-0.910359269	0.028504777	0.27413853	-2.221365629	1.88093E-05	0.001984789	-1.55485324	0.008229491	0.049254962
FBtr0074520	wupA	12.74800803	-0.907632336	0.028936139	0.274893318	-2.219146662	1.89515E-05	0.001984789	-1.552782038	0.008292469	0.049254962
FBtr0074523	wupA	12.74855193	-0.903021804	0.030146161	0.282938062	-2.207226332	2.03829E-05	0.001984789	-1.556004836	0.008241849	0.049254962
FBtr0074521	wupA	12.74790439	-0.898654364	0.030710503	0.284803356	-2.214488102	1.97396E-05	0.001984789	-1.549397944	0.008450519	0.049254962
FBtr0308237	wupA	12.7437449	-0.891296455	0.032344217	0.291826879	-2.235166546	1.7044E-05	0.001984789	-1.562432679	0.007971102	0.049254962
FBtr0074524	wupA	12.76416713	-0.880096549	0.032591705	0.291826879	-2.146630673	3.36564E-05	0.002184859	-1.487447887	0.011321133	0.051131104
FBtr0308236	wupA	12.7442904	-0.886359696	0.033737523	0.298653751	-2.22294878	1.83689E-05	0.001984789	-1.565670413	0.007923843	0.049254962
FBtr0336897	wupA	12.75290102	-0.873919163	0.034872221	0.303133579	-2.192972847	2.35169E-05	0.0020275	-1.515088054	0.009858634	0.049254962
FBtr0308235	wupA	12.75995197	-0.866583041	0.0356563	0.305233598	-2.160456836	3.07678E-05	0.002178922	-1.495674681	0.010951666	0.050184401
FBtr0078477	CG7470	8.81746033	0.697264628	0.047986673	0.370115029	0.386602578	0.390798508	1	1.265890616	0.004885971	0.037684862
FBtr0085635	Fer1HCH	9.363727423	0.675783178	0.055871203	0.406763248	-0.000367216	0.999366708	1	1.694361282	0.000120213	0.002530963
FBtr0085634	Fer1HCH	9.699257381	0.657097589	0.061055567	0.421337632	0.00432461	0.992293507	1	1.731672134	0.000100993	0.002185387
FBtr0074193	CG8952	6.670668586	0.731026919	0.059714609	0.421337632	0.155311654	0.760606801	1	1.548362481	0.001473923	0.020712974
FBtr0330403	RpL37a	7.582436113	0.6788506	0.060318134	0.421337632	0.270225426	0.567561926	1	1.344408958	0.004366221	0.036572966
FBtr0085633	Fer1HCH	9.700386315	0.649766758	0.06397286	0.429891914	0.001549665	0.997234011	1	1.735311839	9.76485E-05	0.002173377

Figure S3.2 (Continued)

FBtr0308333	RpL15	9.632780748	0.650786705	0.064014714	0.429891914	0.222341669	0.614990453	1	1.276070587	0.003428411	0.031893276
FBtr0113742	RpL15	9.632935551	0.653600951	0.062979408	0.429891914	0.224059241	0.61232408	1	1.270493121	0.003575302	0.032766587
FBtr0308334	RpL15	9.612914294	0.641648673	0.068100444	0.447014087	0.193493233	0.661548461	1	1.2605478	0.003816	0.034168555
FBtr0100231	RpL41	8.810529585	0.625740953	0.074631282	0.46552379	0.389596203	0.386163465	1	1.276296508	0.004483067	0.036955418
FBtr0333126	AcCoAS	8.237294871	0.637023606	0.075481687	0.46552379	0.363198994	0.439190702	1	1.299388103	0.004875643	0.037684862
FBtr0085632	Fer1HCH	9.636065348	0.608584452	0.083435622	0.488694357	-0.005864938	0.989619235	1	1.683514547	0.000155657	0.003109157
FBtr0083804	Vha13	7.810319644	0.59884192	0.097928062	0.524579142	0.429998278	0.35714062	1	1.296963804	0.004244654	0.036059737
FBtr0071592	RpL29	6.982023717	0.59424972	0.114325377	0.524579142	0.141874012	0.768817312	1	1.271065499	0.006351318	0.04378475
FBtr0088527	RpL31	8.131072589	0.562134836	0.117845122	0.524579142	0.172084479	0.704280201	1	1.186686095	0.007890406	0.049254962
FBtr0078056	RpLP1	8.66191231	0.578612542	0.104357473	0.524579142	0.011927222	0.978937952	1	1.164095751	0.009017605	0.049254962
FBtr0086273	RpS18	8.366453665	0.578718947	0.101535686	0.524579142	0.129007531	0.778943214	1	1.151859224	0.010541377	0.049254962
FBtr0088525	RpL31	8.234348248	0.53375937	0.135304799	0.550529429	0.245896694	0.587860134	1	1.16657019	0.009193423	0.049254962
FBtr0071135	RpS6	8.812824087	0.493182982	0.159383449	0.600555098	0.144536022	0.74812387	1	1.238980342	0.006215074	0.04378475
FBtr0072188	tsr	6.846540883	0.518061565	0.17585316	0.626133231	0.471741403	0.327679035	1	1.463171199	0.002176436	0.027848726
FBtr0088035	Ef1alpha48D	11.31582917	0.467785091	0.176997993	0.626133231	0.145307202	0.74266572	1	1.153153587	0.008886954	0.049254962
FBtr0072173	eIF-5A	8.912274596	0.44606917	0.204259137	0.691816816	0.216390996	0.6309181	1	1.178778822	0.007820974	0.049254962
FBtr0100861	mt:CoI	16.22051494	-0.403953351	0.252343223	0.773918783	-0.143659718	0.738600544	1	-1.241999097	0.005062325	0.038286907
FBtr0100483	Pglym78	8.265840791	-0.396305872	0.274774643	0.77405815	-0.200768114	0.650953735	1	-1.337356046	0.004541398	0.036955418
FBtr0100482	Pglym78	8.269949184	-0.393938269	0.277622039	0.77405815	-0.203902664	0.645912267	1	-1.336999475	0.004554198	0.036955418
FBtr0076593	Prm	9.879726395	-0.396032796	0.262596273	0.77405815	-0.66429827	0.133375299	1	-1.211294561	0.007395432	0.048412114
FBtr0070801	RpL35	8.651293996	0.399132016	0.259911014	0.77405815	0.195951202	0.661553738	1	1.150598385	0.010019652	0.049254962
FBtr0085384	Pglym78	8.279576611	-0.389529929	0.282702686	0.780941107	-0.196951966	0.657128843	1	-1.308225777	0.005399607	0.039994089
FBtr0332820	CG43897	10.33689599	-0.353898038	0.312790228	0.81221196	0.228403387	0.600942174	1	1.126205826	0.010180822	0.049254962
FBtr0076594	Prm	9.738528635	-0.348451167	0.324917326	0.829870811	-0.647706876	0.144027359	1	-1.157569761	0.010512918	0.049254962
FBtr0334482	bt	12.45641559	-0.328833613	0.346669343	0.849916937	-0.076283274	0.860292223	1	-1.157079268	0.009760232	0.049254962
FBtr0089793	bt	12.47298657	-0.328252619	0.347362797	0.849916937	-0.079124103	0.855188323	1	-1.148905522	0.01026816	0.049254962
FBtr0089746	#N/A	11.66477126	-0.32661288	0.35866263	0.854428711	-0.161620778	0.710431153	1	-1.699705952	0.00027634	0.005006258
FBtr0113290	CG5028	8.953459839	-0.334976817	0.356476998	0.854428711	-0.175629669	0.688743096	1	-1.298649095	0.004769379	0.037684862
FBtr0334483	bt	12.48287089	-0.321628869	0.357174345	0.854428711	-0.075609154	0.861550334	1	-1.14710665	0.010389275	0.049254962
FBtr0301340	bt	12.48261721	-0.320888863	0.358267741	0.854428711	-0.077314779	0.858467247	1	-1.146941532	0.010399625	0.049254962
FBtr0333918	Tm1	9.129795573	-0.292712696	0.407871274	0.903507976	-0.525096132	0.240025505	1	-1.236266225	0.006963648	0.04597188
FBtr0301960	Tm1	9.070789879	-0.29500091	0.404799559	0.903507976	-0.559547606	0.210937079	1	-1.193218151	0.009082119	0.049254962
FBtr0089747	Mlc2	11.92959086	-0.282132573	0.425773521	0.911202122	-0.101368056	0.814653312	1	-1.674187193	0.00032878	0.0058209
FBtr0304129	MtnE	6.35837712	0.343916959	0.437446849	0.912907327	-0.301432604	0.570790075	1	2.535307966	1.34737E-07	5.83112E-06
FBtr0089566	Zasp66	7.409028894	-0.285610891	0.438179199	0.912907327	-0.077984487	0.86383126	1	-1.368721159	0.004838746	0.037684862
FBtr0307492	Mhc	13.06673328	-0.244375869	0.486091162	0.915850486	-0.306083127	0.484940105	1	-1.353007199	0.003065694	0.031477757
FBtr0307495	Mhc	13.08371155	-0.245836439	0.483537626	0.915850486	-0.30814459	0.482056293	1	-1.34987613	0.003126702	0.031477757
FBtr0307496	Mhc	13.09034573	-0.247653001	0.480375988	0.915850486	-0.309623773	0.479923296	1	-1.348726642	0.003152877	0.031477757
FBtr0301827	Mhc	13.09116984	-0.244867853	0.485341231	0.915850486	-0.308395513	0.481751304	1	-1.348845301	0.00315553	0.031477757
FBtr0080896	Mhc	13.08279688	-0.246038838	0.483167333	0.915850486	-0.309095642	0.480693532	1	-1.347925447	0.003166034	0.031477757
FBtr0080895	Mhc	13.06875776	-0.244767428	0.485404578	0.915850486	-0.306608108	0.484210074	1	-1.348120868	0.003166146	0.031477757
FBtr0080901	Mhc	13.08201556	-0.247573292	0.480428488	0.915850486	-0.309971254	0.479363714	1	-1.347172048	0.003177043	0.031477757
FBtr0080902	Mhc	13.06443362	-0.239641549	0.494613785	0.915850486	-0.299899043	0.493732749	1	-1.347658218	0.003178655	0.031477757
FBtr0080899	Mhc	13.05021397	-0.238288068	0.497028269	0.915850486	-0.297303511	0.497456435	1	-1.347859689	0.00317881	0.031477757
FBtr0307493	Mhc	13.07839279	-0.250264624	0.475764163	0.915850486	-0.3072163	0.483363099	1	-1.346485455	0.003197893	0.031477757
FBtr0080900	Mhc	13.0835511	-0.248521696	0.478740285	0.915850486	-0.310637564	0.478449192	1	-1.346045252	0.003200345	0.031477757
FBtr0080898	Mhc	13.06951931	-0.247269841	0.480931192	0.915850486	-0.308160716	0.481941695	1	-1.346223786	0.003200749	0.031477757
FBtr0301828	Mhc	13.06768918	-0.244690978	0.485530151	0.915850486	-0.305578598	0.485659228	1	-1.345885012	0.003211468	0.031477757
FBtr0080903	Mhc	13.06519742	-0.24216213	0.490063034	0.915850486	-0.301474568	0.491407359	1	-1.34574732	0.003213604	0.031477757
FBtr0307494	Mhc	13.06572596	-0.236338386	0.500837365	0.915850486	-0.294179769	0.50217685	1	-1.346138671	0.003232632	0.031477757
FBtr0080907	Mhc	13.13336714	-0.237879213	0.498383892	0.915850486	-0.293211625	0.503825629	1	-1.342079917	0.003330633	0.031893276
FBtr0301829	Mhc	13.11204079	-0.24022667	0.49404088	0.915850486	-0.295884298	0.499844384	1	-1.337109317	0.00342776	0.031893276
FBtr0080897	Mhc	13.10379567	-0.241377489	0.491885517	0.915850486	-0.296566113	0.498789394	1	-1.336188321	0.00343907	0.031893276
FBtr0089967	Tm1	9.292274613	-0.239418305	0.49724263	0.915850486	-0.516248329	0.246423656	1	-1.183229062	0.009484203	0.049254962
FBtr0082158	MtnA	9.216561836	0.262046542	0.515638327	0.923347903	-0.336238006	0.47457409	1	1.898711031	4.39015E-05	0.001103202
FBtr0080905	Mhc	13.14242867	-0.227860678	0.516809725	0.923347903	-0.277640006	0.526642588	1	-1.349024027	0.003210859	0.031477757
FBtr0084901	CG5028	8.68797227	-0.238616102	0.511531227	0.923347903	-0.224542797	0.612231139	1	-1.286982764	0.005624378	0.040947571
FBtr0301959	Tm1	9.30108322	-0.23282025	0.509109144	0.923347903	-0.506225627	0.255445976	1	-1.18141516	0.009601266	0.049254962
FBtr0080906	Mhc	13.14367919	-0.221590464	0.528328561	0.931882701	-0.2743475	0.53155027	1	-1.351343369	0.003165865	0.031477757
FBtr0080167	l(2)06225	11.00621224	0.201585384	0.592495476	0.936987748	-0.018713753	0.965373381	1	3.045750402	1.52801E-11	2.38064E-09
FBtr0074910	fln	9.093107016	0.205253396	0.58783586	0.936987748	0.418882319	0.363308236	1	-1.686222144	0.000860214	0.01313935
FBtr0083030	Mf	9.347502964	-0.219177136	0.544708685	0.936987748	-0.170449194	0.696720071	1	-1.429261798	0.001777759	0.023877145
FBtr0337053	Hsc70-4	9.925609164	0.209758762	0.548732352	0.936987748	-0.056991918	0.900690215	1	1.175358624	0.009214622	0.049254962
FBtr0083059	Hsc70-4	9.923502261	0.210795056	0.546574147	0.936987748	-0.068555974	0.880721584	1	1.175589365	0.009222994	0.049254962
FBtr0083055	Hsc70-4	9.920834788	0.210900111	0.546541794	0.936987748	-0.067169491	0.883089428	1	1.173986133	0.009314814	0.049254962
FBtr0083058	Hsc70-4	9.920834788	0.210900111	0.546541794	0.936987748	-0.067169491	0.883089428	1	1.173986133	0.009314814	0.049254962
FBtr0083060	Hsc70-4	9.921259102	0.209814279	0.548611271	0.936987748	-0.067575823	0.88239267	1	1.173578302	0.009343053	0.049254962
FBtr0083056	Hsc70-4	9.923511694	0.207509173	0.552963012	0.936987748	-0.073034714	0.872959303	1	1.170280961	0.009523847	0.049254962
FBtr0301923	up	10.09954103	-0.191989741	0.58551652	0.936987748	-0.282763456	0.517364173	1	-1.167693782	0.01025091	0.049254962
FBtr0100563	up	10.10029312	-0.197031381	0.575760428	0.936987748	-0.277121688	0.525685579	1	-1.162703776	0.01054371	0.049254962
FBtr0308234	up	10.10307782	-0.194846297	0.580035672	0.936987748	-0.274138826	0.530073494	1	-1.162434981	0.010559151	0.049254962
FBtr0332723	fln	9.146086756	0.196124294	0.604418857	0.93980497	0.379148708	0.409256826	1	-1.685907602	0.000832183	0.012965404
FBtr0089562	Zasp66	7.505697387	-0.186105009	0.612423084	0.944956098	-0.039457151	0.930531993	1	-1.231281089	0.010451857	0.049254962
FBtr0308683	CoVIb	8.881257995	-0.172201811	0.630315093	0.962775406	0.034484249	0.937582603	1	-1.198624725	0.009445603	0.049254962
FBtr0074815	CoVIb	8.951532218	-0.165959932	0.642335824	0.965983797	0.032871353	0.940375089	1	-1.206762547	0.008975503	0.049254962

Figure S3.2 (Continued)

FBtr0301921	up	9.998410587	-0.159988685	0.649693406	0.966646976	-0.256558725	0.556776555	1	-1.176060015	0.00997931	0.049254962
FBtr0081030	Arr1	8.841354281	-0.166546183	0.647075148	0.966646976	-1.603857052	0.000883754	0.028685175	1.04136945	0.020567972	0.074523023
FBtr0308233	up	10.02173147	-0.153714549	0.662333491	0.97167192	-0.236124634	0.588756348	1	-1.192758946	0.009032142	0.049254962
FBtr0301919	up	9.999304201	-0.154419966	0.661100958	0.97167192	-0.260255455	0.551157269	1	-1.171969895	0.010240526	0.049254962
FBtr0110872	l(2)06225	11.0336913	0.142849674	0.702837599	0.98094734	0.01260226	0.976795762	1	2.945746235	6.34634E-11	6.17975E-09
FBtr0310535	CG14125	9.868059748	0.148564747	0.70732818	0.98094734	-1.917352437	0.000253803	0.012145604	0.07016217	0.883687629	0.902218431
FBtr0076044	CG14125	10.03070582	0.152791457	0.700194653	0.98094734	-1.918761182	0.000265052	0.012145604	0.063817408	0.89445325	0.909460874
FBtr0083032	Mf	8.869792383	-0.129824566	0.716546989	0.987947088	-0.095975174	0.827417436	1	-1.170240535	0.010490947	0.049254962
FBtr0083143	Act88F	10.978493	-0.088103911	0.805263962	0.988849725	-0.019130924	0.964745059	1	-1.676224604	0.000396055	0.006707102
FBtr0089630	CG10910	9.123099807	0.132453313	0.732534499	0.988849725	0.156314552	0.745269671	1	1.418434392	0.003217256	0.031477757
FBtr0077144	CG4769	9.530115143	-0.071435232	0.843580737	0.988849725	-0.083898738	0.846166644	1	-1.326179463	0.003704073	0.033552007
FBtr0085196	Mlc1	8.533053555	0.030101108	0.935355787	0.988849725	0.004233593	0.992301353	1	-1.377245253	0.003896117	0.034489491
FBtr0306086	CG4769	9.322618119	-0.058425019	0.872125978	0.988849725	-0.08287207	0.848038119	1	-1.319070164	0.003941215	0.034496701
FBtr0086303	CG9090	9.400383022	0.020173457	0.955344381	0.988849725	0.271627734	0.532484276	1	-1.335031974	0.004110634	0.03557982
FBtr0305122	CG1746	11.85370627	0.043738975	0.904060587	0.988849725	0.294452816	0.502440044	1	-1.314506351	0.005022785	0.038286907
FBtr0084994	Ald	11.4081457	-0.109129656	0.760274063	0.988849725	0.066973948	0.877262223	1	-1.278846606	0.005160008	0.038650441
FBtr0085774	CG1746	11.92184796	0.056988485	0.875188218	0.988849725	0.292604975	0.505591702	1	-1.301851839	0.005442071	0.039994089
FBtr0085772	CG1746	11.95873861	0.059521726	0.869768377	0.988849725	0.292349908	0.506268268	1	-1.287833313	0.005974527	0.043094044
FBtr0083857	ninaE	10.70726439	0.06356164	0.86259996	0.988849725	-1.218166634	0.007095975	0.178314979	1.165292171	0.007550181	0.048608188
FBtr0331936	CG7712	7.125968058	-0.01853489	0.960315173	0.988849725	0.258812423	0.56899981	1	-1.326557605	0.007505124	0.048608188
FBtr0306237	CG1746	10.94662408	0.061426473	0.866807949	0.988849725	0.161805148	0.720658561	1	-1.230839361	0.009006215	0.049254962
FBtr0085195	Mlc1	8.517058684	0.002702449	0.994225799	0.999524566	0.020428117	0.963633435	1	-1.419099573	0.002997198	0.031477757
FBtr0084892	Npl4	9.245821212	-0.000683439	0.998895946	0.999524566	0.516593856	0.229382463	1	-1.247732438	0.006098485	0.043584589
FBtr0300730	Npl4	9.234933824	0.002310775	0.99457009	0.999524566	0.517095432	0.22901997	1	-1.24201829	0.006322166	0.04378475

Differential expression analysis for HECTs was conducted using the edgeR Bioconductor software package.(*2*)

Literature Cited

1. M. E. Ritchie *et al.*, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* (2015), doi:10.1093/nar/gkv007.

2. M. Robinson, D. McCarthy, G. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

3. M. D. Paraskevopoulou *et al.*, DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Research* **41**, W169–73 (2013).

4. H. Mi *et al.*, The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Research* **33**, D284–8 (2005).